THE UNIVERSITY *of* EDINBURGH

Post-Translational Regulation of Nanog and Nanog-Interacting Proteins in Mouse Embryonic Stem Cells

Marcia Roy

Thesis submitted for the degree of Doctor of Philosophy

Institute of Cell and Molecular Biology

Wellcome Trust Centre for Cell Biology

University of Edinburgh

2012

DECLARATION

I declare that all work presented in this thesis is my own unless otherwise stated. Furthermore, this work has not been submitted for any other degree or professional qualification.

Marcia Roy

2012

ACKNOWLEDGEMENTS

Firstly, I would like to thank my first supervisor Mike Tyers who made the bold decision to move to the University of Edinburgh and establish a laboratory equipped with state-of-the-art equipment and encouraged me to have ideas and play with them. This allowed me to grow from the technician I was at the beginning of the PhD studies into an independent, free-thinking and, at times, a prototypical/stereotypical mad scientist. For this I am truly grateful, for if our supervisors simply spoon-fed us for 3-4 years, how would we ever learn to feed ourselves? The work ethic and thinking skills learned in his lab will undoubtedly provide a solid foundation upon which to build my future career in science. Thank you Mike for supporting even my craziest ideas!

Also from the Tyers lab in Edinburgh, I would like to give an enormous thank you to Larissa Christian whose work ethic and dedication ensured that the proteomic data ascertained throughout my PhD was of high quality, reproducible and rapidly obtained. I am forever grateful to Larissa for the large number of gel slices that she faithfully cut, digested and loaded onto the Orbitrap for two years. Her efforts will be appreciated not only by myself but hopefully also by the scientific community as a whole in future publications resulting from this work.

From the Kinetic Parameter Facility (KPF) and the Centre for Systems Biology at Edinburgh (CSBE), I give a tremendous amount of thanks to Dr. Thierry Le Bihan whose expertise in mass spectrometry allowed for such a study to be performed. In addition to providing access to his LTQ-Orbitrap mass spectrometer, he graciously performed the MaxQuant data analysis on all MS data acquired over the past 3-4 years. A special thanks should be given to Thierry for providing me with

the critical/analytical left side of his brain during those times when the overly-creative right side of my brain had taken over.

An enormous thank you to everyone from the ISCR who assisted me throughout the course of my PhD. To Sally Lowell for providing me with my first vial of mouse embryonic stem cells. To Ian Chambers for taking me under his wing, providing reagents and cell lines and trying to keep me focused by repeatedly asking me "Tell me something new about Nanog!". To Dougie Colby who trained me how to culture ES cells as per the Chambers' lab's protocols.

To my daughter Clara Jeanne (CJ) who reminds me on a daily basis why I work so hard, what really matters at the end of the day and that eating chocolate usually makes everything all better again.

I thank my family for their continued support and compassion over the past 4 years. Their encouragement and belief in my abilities helped me to get through those periods when my PhD seemed to be an enormous and famished monster who could never be satisfied no matter how much data I threw at it. Moreover, they serve as a constant reminder that honest, hard-working, law-abiding people actually do exist in this world. Their honesty and demonstration of basic human decency on a daily basis prevented me from losing my faith in humanity during my final 10 months in the Tyers lab in Toronto as well as throughout my time at the University of Edinburgh.

Lastly, I would like to thank those voices in my head who "although I know they're not real, they have some pretty good ideas sometimes." *Happy Bunny*

This work was supported by a University of Edinburgh SBS studentship and by funding from the Scottish Universities Life Sciences Alliance (SULSA).

ABSTRACT

Pluripotent embryonic stem cells (ESCs) possess an unlimited capacity for self-renewal. This property of ES cells is both defining and unique. Harnessing this potential of ESCs would provide tremendous opportunity in the field of regenerative medicine and its attempts to combat degenerative diseases such as Parkinson's, muscular dystrophy, etc.

In 2006, Shinya Yamanaka was able to demonstrate that the ectopic expression of four proteins could reverse the process of differentiation and provide somatic cells with the characteristics ESCs. One year later, James Thompson's group proved the same feat could be accomplished in human somatic cells using a different set of four proteins, including Nanog. The prospect of converting one's own cells into a stem cell which could subsequently differentiate and repopulate an area of the body afflicted by gross degeneration was revolutionary. In the years following Yamanaka's and Thompson's discoveries, however, there has been little insight gained into how these proteins are regulated post-translationally.

In this study, four proteins which had previously been identified by Yamanaka as being 'pluripotency factors' were used as baits in order to ascertain a protein-protein interaction network. This network was subsequently interrogated using various chemical compounds and small molecules in order to dissect the signal transduction pathways feeding into pluripotency, as well as, post-translational modifications regulating the factors themselves.

In this way, the chemical inhibitor H89 was found to decrease the presence of Nanog phosphorylation and possibly its dimerization resulting in the Nanog protein being destabilized and targeted for degradation. Inversely, the pan-cullin inhibitor

MLN4924 was identified to increase the abundance of both phosphorylated Nanog and total Nanog protein. In an attempt to identify the Cullin Ring Ligase (CRL) responsible for the degradation of Nanog protein in ESCs, each cullin identified in the protein interaction network was inhibited using specific shRNAs. Quantitative fluorescence microscopy was performed and identified that inhibition of CUL3 increases Nanog protein levels, suggesting that a CUL3-based CRL may be responsible for the post-translation regulation of Nanog.

Additionally, the quantitation of Sox2 protein levels in CUL4B shRNA cell line demonstrates that Sox2 protein levels may be regulated by a CUL4B-based CRL. Further studies will reveal whether or not CUL4A depletion also results in elevated Sox2 protein levels. If not, this would include the pluripotency factor Sox2 among the recently identified CUL4B-isoform-specific substrates for degradation and possibly provide the basis for a hypothesis of developmentally regulated substrate specificity.

In addition to MLN4924, several other small molecules were identified as being able to increase phospho-Nanog protein levels in this study. Among them were the cell permeable peptides Ht-31 and PKI (14-22) amide. These peptides were found to both stabilize phospho-Nanog and produce ES cell colonies that uniformly express the Nanog protein. The development of a growth medium containing these peptides in order to maintain homogeneous pluripotent ES cells is currently in progress and received backing for a patent application by the University of Edinburgh on February 23, 2012.

TABLE OF CONTENTS viii-xvi

APPENDICES:

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BN-PAGE	Blue Native polyacrylamide gel electrophoresis
BSA	Bovine serum albumin
CTRL	Control
DAPI	4', 6-Diamidino-2-phenylindole dihydrochloride
DMSO	Dimethylsulfoxide
ddH_2O	De-ionized, distilled water
D-PBS	Dulbecco's phosphate-buffered saline
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium bromide
ESCs	Embryonic stem cells
ES	Embryonic stem
FBS	Fetal bovine serum
x g	x gravity
g	Gram
GMEM	Glasgow minimal essential medium
H_2O	Water
HU	Hydroxyurea
ID	Identification/Identifier
IP	Immunopurification
IP-LC-MS/MS	Immunopurification-liquid chromatography-Mass spectrometry/Mass spectrometry
kDa	kiloDalton
L	Liter
LB	Luria-Burtani
LC-MS/MS	liquid chromatography-Mass spectrometry/Mass spectrometry
LIF	Leukemia inhibitory factor
M	Molar
MEM	Minimal essential medium
MetOH	Methanol
mESCs	mouse embryonic stem cells
mg	Milligram
$MgCl_2$	Magnesium chloride
mL	Milliliter
mM	Millimolar
MOPS	3-[N-morpholino] propanesulfonic acid
MS/MS	Mass spectrometry/Mass spectrometry
MW	Molecular weight
NaCl	Sodium chloride
$NaHCO_3$	Sodium bicarbonate

ng	Nanogram
nM	Nanomolar
NaOH	Sodium hydroxide
NEAAs	Non-essential amino acids
OD	Optical density
PBS	Phosphate-buffered saline
rpm	Revolutions per minute
RT	Room temperature
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
TBS	Tris-buffered saline
TCA	Trichloroacetic acid
Tris	Hydroxymethyl amino methane
Triton® X-100	Polyethylene glycol *tert*-octylphenyl ether
Tween20	Polyoxyethylene sorbitan monolaurate
µL	Microlitre
µM	Micromolar
UT	Untreated
V	Volts
WT	Wild-type

CHAPTER 1: Introduction

The body of work presented in this thesis encompasses a wide range of topics. As a result, a review of the literature for each topic will be presented in this Chapter. In the first section (1.1), I will describe the developmental stage from which pluripotent mouse embryonic stem cells are derived a well as their maintenance in culture. In the second section (1.2), I will present the discovery of pluripotency factors in mouse embryonic stem cells and their ability to confer the pluripotent embryonic stem cell phenotype. In the third section (1.3), I will provide an update on protein-protein interaction networks for pluripotency factors obtained using an IP-LC-MS/MS strategy. In the fourth section (1.4), I will detail the post-translational regulation of protein abundance in cells by the ubiquitin proteasome system. In the fifth section (1.5), I will present the Cullin-RING ligase family of E3 ubiquitin ligases and the mechanism by which they target their specific substrates for degradation by the 26S proteasome. In the sixth section (1.6), I will provide an overview of chemical compounds and small molecules that have been discovered to provide a control over the various cell fate decisions made by pluripotent stem cells.

1.1 The derivation and culture of mouse embryonic stem cells

The initial stages of mammalian reproduction involve the establishment of a fertilized egg, that is, the fusion of an oocyte and a sperm cell. This cell is now classified as a totipotent stem cell. Totipotent is the term assigned to cells that can give rise to all the different types of cells in the body and construct an entire organism (Mitalipov and Wolf, 2009). In the subsequent rounds of cell division, cells

maintain their totipotency until they reach the blastocyst stage of development (Do *et al.*, 2007). The blastocyst is an early stage embryo comprised of approximately 50 to 100 cells and containing a cluster of cells called the inner cell mass (ICM).

These embryonic stem cells in the ICM are pluripotent and have an apparently unlimited capacity for self-renewal (Suda *et al.*, 1987). Pluripotent is the term assigned to cells that are derived from the ICM of the early developing embryo (blastocyst) and give rise to any cell type found in the three primary germ layers of the embryo (endoderm, mesoderm and ectoderm), as well as germ cells (Beddington and Robertson, 1989; Gardner and Brook, 1997; NIH – Report on Stem Cells, 2001; Nichols, 2001; Hubner *et al.*, 2003; Toyooka *et al.*, 2003). Unlike totipotent cells, pluripotent stem cells do not contribute to the formation of the placenta or extra-embryonic membranes.

At this stage of embryonic development, the previously totipotent stem cells have differentiated and formed a layer of trophoblast cells enveloping a mass of embryonic stem cells. An important function of the trophoectoderm layer of the blastocyst-stage embryo is the secretion of the cytokine Leukemia Inhibitory Factor (LIF). The stem cells residing within the ICM widely express the LIF receptor and rely heavily on LIF-dependent signaling in order to maintain their pluripotency through the activation JAK/STAT and MAPK signalling cascades (Burdon *et al.*, 1999; Niwa *et al.*, 1998; Smith 2001).

In 1981, Evans and Kaufman became the first to report the establishment of mouse pluripotent stem cells in culture (Evans and Kaufman, 1981). Figure 1.1 illustrates the developmental stage from which pluripotent stem cells are derived as well as their potential to form all three germ layers of the body.

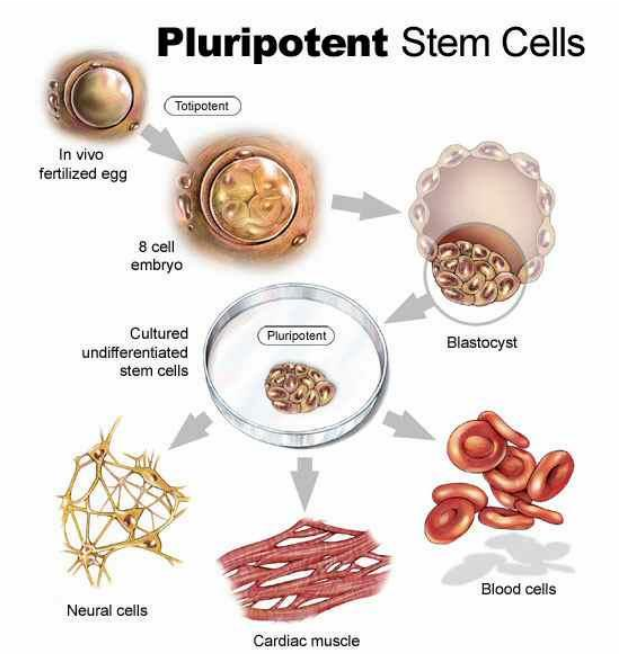


Figure 1.1: Illustration of the origin and potential of pluripotent stem cells. Image taken from www.stemcellresearchfoundation.org.

As a consequence of removing stem cells from the ICM and placing them in culture, however, is that they are deprived of the supply of LIF that the trophoectoderm layer naturally provides. Therefore, stem cells were required to be maintained on primary mouse embryonic fibroblasts (MEFs) which would secrete LIF into the culture medium and inhibit differentiation. The identification of LIF's ability to maintain mouse embryonic stem cells' pluripotential in culture enabled the culturing the cells in feeder-free conditions (Smith *et al.*, 1988; Williams *et al.*, 1988; Peace and Williams, 1990; Nichols *et al.*, 1990). Subsequent modifications and improvements in culture techniques have made the routine maintenance and differentiation of pluripotent mouse embryonic stem cells in monolayer culture possible (Smith *et al.*, 1992; Nichols *et al.*, 2006).

The ability to maintain pluripotent stem cells as monolayers in culture provided several advantages for fundamental research. The main advantage was the

removal of the MEF feeder layer which made embryonic stem cell cultures a heterogeneous population of different cell types. As a result, it was difficult to use stem cell cultures to perform certain types of experimental studies, i.e., quantitative mass spectrometry, IP-LC-MS/MS, to name a few.

Another advantage is that the growth of a uniform monolayer of stem cells made reliable small molecule screening studies possible. In a monolayer culture, each cell is exposed to exactly the same concentration of compound for exactly the same amount of time thereby ensuring a uniform response by the cells and less variation between the conditions being tested. In the case of chemically inducing stem cells to differentiate into a particular cell type, for example, a monolayer culture system ensures that the required differentiation transcriptional programs are activated uniformly and in the maximum number of cells. This optimized level of control over cell fate decisions is simply not possible in a traditional spheroid culture system.

In addition, a monolayer culture system allows for the individual responses of each cell in the population to a particular perturbation to be accurately visualized and quantified using fluorescence-based confocal microscopy assays. Thus, enabling highly sensitive measuring and monitoring of the various cell fate decisions being made by cells throughout entire populations of stem cells.

1.2 Pluripotency factors in mouse embryonic stem cells

In addition to their ability to generate all tissue-specific cell types of the body, embryonic stem cells can also reprogram differentiated cells when they are fused to a somatic cell (Tada *et al.*, 2001). This observation made by Tada *et al.*, 2001, indicated that embryonic stem cells contain unknown factors that can confer pluripotency to somatic cells and drew the interest of researchers in the field of

regenerative medicine. Among them, was Shinya Yamanaka from the Institute for Frontier Medical Sciences at the University of Kyoto in Japan who has contributed significantly to this field (Mitsui *et al.*, 2003; Maruyama *et al.*, 2005; Takahashi *et al.*, 2006; Takahashi *et al.*, 2007; Okita *et al.*, 2007). After having characterized several genes which are specifically expressed in embryonic stem cells, Yamanaka selected 24 of these genes for further testing of their ability to confer pluripotency to mouse embryonic fibroblasts (Takahashi *et al.*, 2006).

Takahashi and Yamanaka reported the ability to reprogram mouse fibroblasts via retroviral transduction of four transcription factors, Oct4, Sox2, KLF4 and cMyc (Takahashi and Yamanaka, 2006). The following year, they demonstrated the ability to reprogram adult human skin cells using the same four defined factors (Takahashi *et al.*, 2007). Another group reported the ability of a different combination of four factors, (Nanog, Oct4, Sox2 and Lin28) were sufficient to reprogram human foreskin fibroblasts to pluripotent stem cells, (Yu *et al.*, 2007).

The study by Yu *et al*., 2007 revealed that the four factors used by Yamanaka were not exclusive. This led to the relentless pursuit of undiscovered pluripotency-inducing factors in an attempt to further define the exact composition of the mysterious cocktail of factors contained within the stem cell nucleus that was first described by Tada *et al*., 2001 more than a decade ago.

A summary of the various methods routinely used by scientists to generate induced pluripotemt stem cells (iPSCs) from somatic cells using Yamanaka's four factors is illustrated in Figure 1.2.

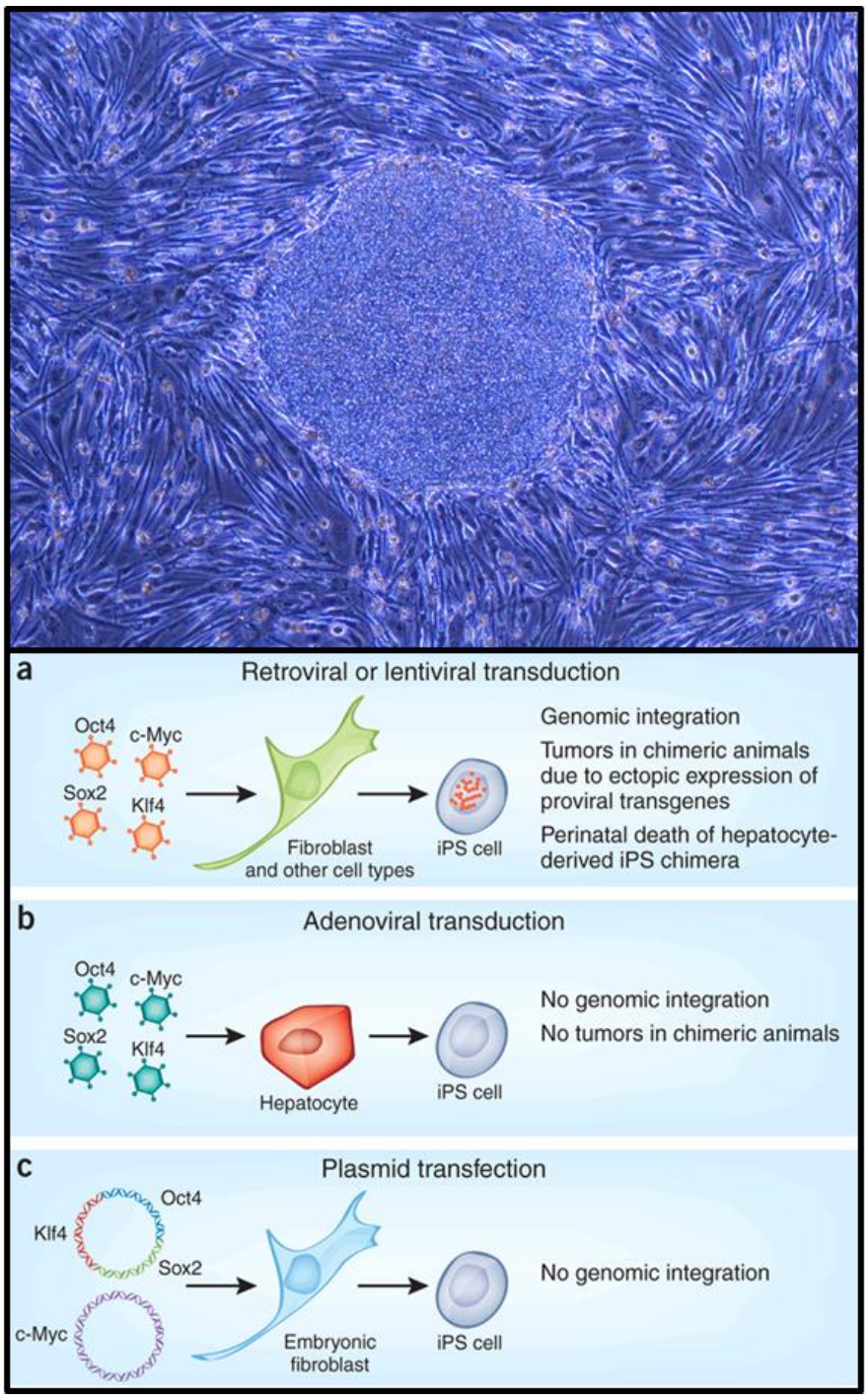


Figure 1.2: Methods of generating induced pluripotent stem cells (iPS).
Image in TOP panel taken from Okita *et al.*, 2007. Image in the BOTTOM panel taken from Lowry and Plath, 2008.

In the years that have followed these landmark publications, multiple groups have reported the ability to reprogram mouse and human somatic cells using

different factors such as Esrrβ, Nr5a2 (Feng *et al.*, 2009; Heng *et al.*, 2010), different numbers of factors such as Sox2 and Oct, Oct4 alone (Huangfu *et al.*, 2008; Kim *et al.*, 2009), using a plasmids (Okita *et al.*, 2008; Stadtfeld *et al.*, 2008; Woltjen *et al.*, 2009; Kaji *et al.*, 2009), in combination with small molecules (Shi *et al.*, 2008; Maherali *et al.*, 2009) and by direct delivery of the proteins themselves using defined factors (Li *et al.*, 2009; Kim *et al.*, 2009). These rapid developments are the result of the intense pursuit of patient-specific iPSCs for the treatment of degenerative diseases such as Parkinson's or Alzheimer's.

The work of Yamanaka has resulted in the establishment of an entirely new branch of biology. More importantly, it has challenged our fundamental beliefs about mammalian development which has always been thought to occur linearly and in one direction (Beddington and Robertson, 1989; Gardner and Brook, 1997). We now know that this process can be reversed and we have refined our definition of a pluripotency factor. A pluripotency factor can currently be defined as any protein that can maintain the pluripotency of an embryonic stem cell and can also confer pluripotency to a somatic cell.

1.3 Protein-protein interaction mapping using LC-MS/MS

While much work has been done to characterize the genes either expressed exclusively or upregulated in embryonic stem cells, very few datasets have been published with regards to the protein-protein interactions in which pluripotency factors are involved. Exactly how Yamanaka's four reprogramming factor proteins confer pluripotency to somatic cells remains somewhat elusive. Considering that the majority of pluripotency factors are transcription factors, it is generally assumed that they confer pluripotency to somatic cells by activating the transcription of other

embryonic stem cell-specific genes. In this way, they change both the transcriptional and proteomic composition of the cell from a tissue specific mixture to one that is characteristic of a pluripotent cell. This view of reprogramming may be a gross over-simplification especially when considering that one of the initial factors reported as being capable of reprogramming human cells was not a transcription factor at all, i.e., the mRNA binding protein Lin28 (Yu *et al.*, 2007). Two of the more recently identified reprogramming factors are the hormone responsive transcriptional repressors and co-activators Esrrβ and Nr5a2 (Feng *et al.*, 2009; Heng *et al.*, 2010). These reports suggest that the transcriptional repression of certain sets of genes may be equally as important as the activation of the Oct4, Sox2 and Nanog core transcriptional circuitry in ES cells. Furthermore, these factors suggest that the transcriptional activation and repression mechanisms required for the induction of pluripotency are further integrated with extracellular signaling.

One approach to deciphering the mechanisms by which reprogramming occurs is to understand the protein-protein interactions of these factors. Identifying the protein interaction networks of key pluripotency factors such as Nanog would provide a view of the Nanog-imposed stem cell state beyond its ability to bind to DNA. Furthermore, identification of the protein interaction networks of multiple pluripotency factors provides insight as to whether or not these networks converge or if they are unique. These networks may possibly reveal a much larger and more complex set of inter-dependent signalling pathways that govern pluripotency.

Since 2006, the view of pluripotency factors and their contribution to the pluripotent stem cell state has been primarily restricted to the nucleus and the transcriptome (Chambers, 2004; Boyer *et al.*, 2005; Ivanova *et al.*, 2006; Loh *et al.*,

2006; Matoba *et al.*, 2006; Babaie *et al.*, 2007; Kim *et al.*, 2008). By comparison, there have been only two published reports of protein-protein interaction mapping studies using multiple baits in mouse embryonic stem cells (Wang *et al.*, 2006 and Van den Berg *et al.*, 2010). The first was the interaction network ascertained for Nanog and four other factors by the Orkin lab at Harvard University (Wang *et al.*, 2006). Using immunopurification of epitope tagged pluripotency factor proteins followed by liquid chromatography coupled with tandem mass spectrometry (IP-LC/MS/MS), a protein interaction network dubbed 'The Roadmap to Pluripotency' was created (Wang and Orkin, 2008). This 'roadmap' would shape stem cell researchers' views about the protein interactions governing pluripotency and reprogramming in the years to come.

The second IP-LC/MS/MS-based approach to mapping pluripotency factor interactions was published by Van den Berg in 2010. Unlike Orkin's Nanog-centered interaction network, Van den Berg and colleagues built an interaction network centered around another master regulator of pluripotency, Oct4. In this publication, epitope-tagged pluripotency factors were stably expressed from a plasmid and nuclear extracts were used as the starting material for all purifications. This very different snapshot of pluripotency factor interactions provided a much broader view of the pluriproteomic landscape than its predecessor.

Recently, however, a very different view of the protein-protein interaction landscape for Oct4 was obtained by using whole cell lysates as the starting material for IP-LC-MS/MS experiments (Pardo *et al.*, 2010). In this study, the CUL4B protein was reported as an interactor with the Oct4 protein interaction network. Although not highlighted by the authors of the study, this marked the first time that an essential

component of a Cullin-based E3 ligase was identified as interacting with a master pluripotency factor. Since Cullin RING E3 Ligases (CRLs) are key complexes involved in the ubiquitination of specific substrate proteins, the identification of a cullin protein with Oct4 implies that post-transcriptional mechanisms are involved in the maintenance of pluripotency in ES cells.

One year later, another study would identify CUL4B as being part of the Oct4 protein-protein interaction network in nuclear lysates (Cheong *et al.*, 2011). The authors of this study concluded that the Oct4-CUL4B interaction identified by IP-LC-MS/MS in the nucleus of mESCs may be due to both proteins independently interacting with beta-catenin (Takao *et al.*, 2006; Tripathi *et al.*, 2007).

The reported protein interaction landscape for Sox2 in mESCs undergoing differentiation (Mallanna *et al.*, 2010) revealed that there are several proteins interacting with both Sox2 and other master pluripotency factors Nanog, Oct4 and Esrrβ (Wang *et al.*, 2006; Liang *et al.*, 2008; Van den Berg *et al.*, 2010; Pardo *et al.*, 2010 and Van den Berg *et al.*, 2008, respectively).

Figure 1.3 shows a Cytoscape diagram integrating all of the IP-LC-MS/MS protein-protein interaction networks reported in the literature to date for 10 different pluripotency factors in mESCs.

As highlighted by Mallanna and Rizzino (2012) and shown in Figure 1.3, the shared interactors identified between these key pluripotency factors implies a high level of connectivity between the 10 pluripotency factors.

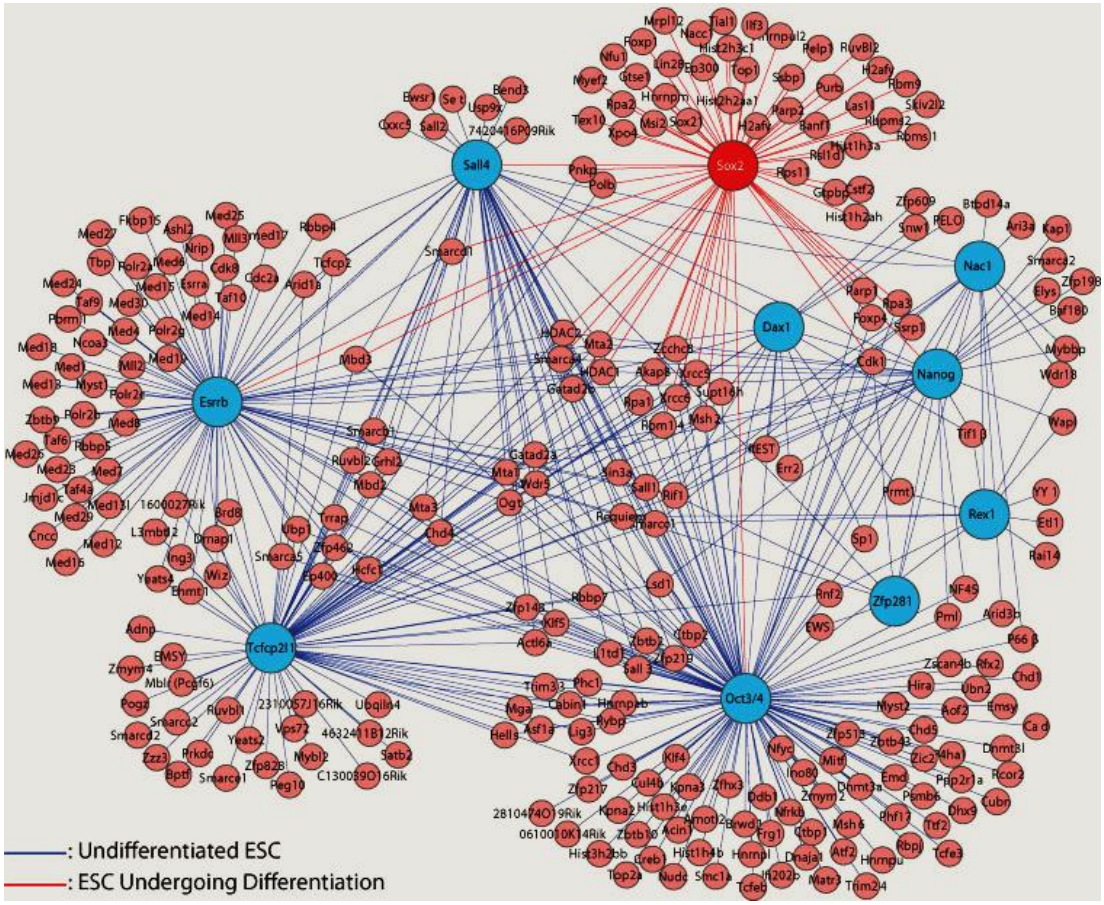


Figure 1.3: The published protein–protein interaction datasets for 10 different pluripotency factors in mESCs. Using Cytoscape, the interaction networks for 10 different pluripotency factors were merged together in order to generate an integrated protein interaction network for mESCs. The interactors identified for Nanog, Oct4, Esrrβ, Sall4, Zfp281, Rex1, Dax1, Tcfcp1l1 and Nacc1 in mESCs (coloured in blue) (Wang *et al.*,2006; Liang *et al.*, 2008; Pardo *et al.*, 2010; Van den Berg *et al.*, 2010) were integrated with the interactors identified for Sox2 undergoing early-stage differentiation (coloured in red) (Mallanna *et al.*, 2010). Image taken from Mallanna and Rizzino, *J Cell Physiology* 2012.

Such a highly connected network of protein interactions involved in the maintenance of pluripotency must be tightly regulated. This is particularly true for the master pluripotency factors themselves since it has been documented that a 2-fold increase in either Sox2 or Oct4 results in the differentiation of mESCs rather than the maintenance of pluripotency (Niwa *et al.*, 2000; Chew *et al.*, 2005; Kopp *et al.*, 2008). Furthermore, the successful reversal of the differentiation process is also highly dependent upon a seemingly delicate balance in the levels of pluripotency

factor expression during reprogramming (Eminli *et al.*, 2008; Papapetrou *et al.*, 2009; Cox and Rizzino, 2010). Therefore, the identification of proteins involved in maintaining this balance would provide valuable insight into how pluripotency can be re-acquired by somatic cells as well as how this potential is both maintained and lost in ES cells.

As previously mentioned, the studies by both Pardo *et al.*, 2010 and Cheong *et al.*, 2011 identified CUL4B as an interactor with Oct4. However, neither study investigated the possible role of a CUL4B CRL being involved in the degradation of the Oct4 protein or any other master pluripotency factor within the network. Furthermore, the Cheong *et al.*, 2011 dataset identified proteins involved in nuclear import and export (Kpna2 and Rcc1) as interacting with Oct4. While Oct4 and Kpna2 have been previously reported to interact (Li *et al.*, 2008), the identification of Rcc1 was novel to the Cheong *et al.*, 2011 study. The cargo being imported to and exported from the nucleus of mESCs by these proteins have not been identified and the authors postulated that these proteins may be working together in order to facilitate the import of unidentified proteins involved in the remodeling of chromatin. However, the possibility of these proteins being involved in the import/export of Oct4 itself or other pluripotency factors was surprisingly not raised in that study.

The observed interaction between Oct4, Kpna2 and Rcc1 in nuclear extracts by Cheong *et al.*, 2011 may provide a key insight into the cytoplasmic export observed for another master pluripotency factor Sox2 (Baltus *et al.*, 2009). In the study by Baltus *et al*, 2009, the cytoplasmic export of Sox2 preceded its ubiquitination and degradation by the 26S proteasome. Combined with the fact that the Pardo *et al.*, 2010 study was performed on whole cell lysates and is the only other

report of the CUL4B interaction, it is plausible that this cullin protein may form a CRL E3 ligase that targets its substrates for degradation in the cytoplasm. Despite the shuttling of Sox2 between the nucleus and cytoplasm, however, the Mallanna *et al*., 2010 study performed Sox2 IP-LC-MS/MS experiments using nuclear extracts as the starting material. Taking into account the Baltus *et al*., 2009 study, a very different snapshot of pluripotency factor regulation could be obtained by performing an IP-LC-MS/MS study using whole cell lysates rather than nuclear extracts.

1.4 Post-translational regulation of protein abundance by the Ubiquitin Proteasome System (UPS)

In order to determine a protein's abundance in cells or tissues, the mRNA transcript level is often quantified and used as an estimate of protein levels and a strong correlation between transcription and translation is assumed (Hack, 2004). However, several studies have revealed that in some cases there is little to no correlation between transcript levels and protein levels in cells and tissues (Humphrey-Smith *et al.*, 1997; Gygi *et al.*, 1999; Futcher *et al.*, 1999; Ideker *et al.*, 2001; Chen *et al.*, 2002; Griffin *et al.*, 2002; Washburn *et al.*, 2003; Hack, 2004). Furthermore, one recent study by Helbig *et al.*, 2011 has revealed that in the case of certain metabolic enzymes, an inverse correlation exists. There are many possible explanations for this lack of correlation between transcription and protein expression level. One possible explanation is the rate and the degree to which protein synthesis and degradation (i.e., protein turnover) is occurring within cells. Information about the level of protein degradation in the cell is simply not captured at the transcriptome level of analysis.

Protein homeostasis is regulated by cellular protein degradation pathways (the lysosomal degradation pathway and the ubiquitin-proteasome pathway). In the Ubiquitin Proteasome System (UPS), proteins are modified by the covalent attachment of ubiquitin to a lysine residue through a cascade of reactions carried out by three enzymes (Etlinger *et al.*, 1977; Hershko *et al.*, 1980; Hershko *et al.*, 1983; Ganoth *et al.*, 1988). The first step of this reaction involves the hydrolysis of ATP and subsequently the adenylylation of a ubiquitin molecule by a ubiquitin activating enzyme (referred to as an E1). This now adenylylated, or 'activated,' ubiquitin molecule is then transferred to the cysteine residue within the active site of the E1 along with the adenylylation of a second molecule of ubiquitin (Haas *et al.*, 1982). The activated ubiquitin molecule residing within the active site of the E1 is then transferred to a cysteine residue within the active site of a second enzyme, a ubiquitin-conjugating enzyme (known as an E2). The E2 then collaborates with a ubiquitin-conjugating enzyme (termed an E3 ligase) in order to transfer the 'activated' ubiquitin molecule to the target protein to be degraded. Using this mechanism, a Lys48 linked polyubiquitin chain containing a minimum of four ubiquitin molecules is attached to the substrate protein in order for it to be targeted to the 26S proteasome for degradation (Thrower *et al.*, 2000).

E3 ubiquitin ligases can be divided into three groups based on whether or not they contain a HECT (homologous to E6-AP carboxy terminus), RING (Really Interesting New Gene) finger or a U-box domain containing protein (Ardley and Robinson, 2005). After the formation of a thioester intermediate with ubiquitin, HECT E3 ligases transfer ubiquitin directly to the substrate protein (Kee and Huibregtse, 2007). RING-finger E3 ligases, however, serve as scaffolds bringing the

substrate protein in close proximity with the ubiquitin-charged E2 which transfers ubiquitin directly to the substrate (Willems *et al.*, 2004). The RING-finger E3s can be further sub-divided into three groups: the Cullin-RING E3 ligases (Willems *et al.*, 2004; Petroski and Deshaies, 2005; Deshaies and Joazeiro, 2009), RING-finger E3 ligases in which the RING-finger and substrate binding domains are located within the same protein (Sun, 2003) and the U-box (or RING-like) domain-containing E3s. While the U-box domain is structurally similar to the RING-finger domain, it does not contain the metal-chelating residues which are characteristic features of the RING domain (Hatakeyama and Nakayama, 2003). A schematic of the mechanism by which ubiquitin is conjugated to proteins by the UPS is illustrated in Figure 1.4.

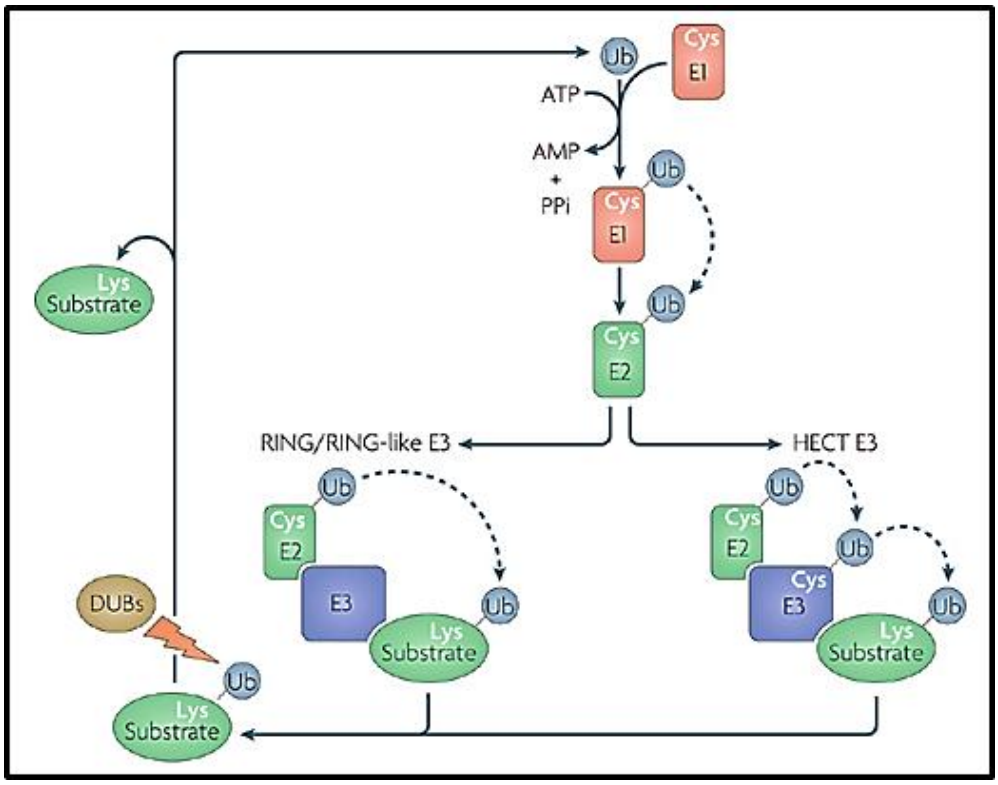


Figure 1.4: The conjugation of ubiquitin to substrate proteins. The diverse ubiquitin conjugation processes performed by the HECT and RING/RING-like (U-box) E3 ligases is shown in this figure. The reversible nature of ubiquitin conjugation by De-ubiquitinating enzymes (DUBs) is also illustrated. Abbreviations: Ub = ubiquitin; Lys = Lysine; Cys = Cysteine; ATP = Adenosine triphosphate; ADP = Adenosine diphosphate; PPi = pyrophosphate. Image taken from Ravid and Hochstrasser (2008).

The covalent attachment of Lys-48 polyubiquitination chains to a protein does not always result in its degradation. Instead, this process may be reversed by a specialized class of enzymes called de-ubiquitinating enzymes (DUBs) (Wilkinson, 1997). DUBs provide an additional level of regulation and control over protein abundance in cells prior to protein degradation by the 26S proteasome (Iyer *et al.*, 2004; Reyes-Turcu *et al.*, 2009). They have several essential roles in the ubiquitin proteasome pathway: processing of ubiquitin precursors, ubiquitin recycling, proofreading of protein ubiquitination and the disassembly of polyubiquitin chains (Amerik and Hochstrasser, 2004; Katz *et al.*, 2010; Burrows and Johnston, 2012).

1.5 Cullin-RING E3 ubiquitin Ligases (CRLs)

The mammalian genome encodes 7cullin proteins: CUL1, CUL2, CUL3, CUL4A, CUL4B, CUL5 and CUL7. Cullins are key components of the multi-subunit complexes known as Cullin RING Ligases (CRLs) responsible for the ubiquitination of proteins and targeting them for degradation (Kipreos *et al.*, 1996; Petroski *et al.*, 2005; Zheng *et al.*, 2002). With the exception of CUL3-based CRLs, all cullin-based E3 RING ligases contain both a substrate recognition subunit (SRS) as well as a substrate-specific adaptor protein in order to bind to their respective targets (Willems *et al.*, 2004; Angers *et al.*, 2006; Kamura *et al.*, 2004). Each of these adaptor proteins specifically binds to their particular cullin through a characteristic cullin-binding motif as illustrated in Figure 1.5. In turn, these substrate-specific adaptor proteins bind directly to the target protein to be degraded (substrate) thereby conferring substrate specificity to the CRL (Craig *et al.*, 1999; Winston *et al.*, 1999; Mahrour *et al.*, 2008; Jackson *et al.*, 2009). In the case of CUL3-based CRLs, a BTB/POZ

domain-containing protein links this cullin directly to its specific target. (Geyer *et al.*, 2003; Xu *et al.*, 2003; Pintard *et al.*, 2003; Pintard *et al.*, 2004).

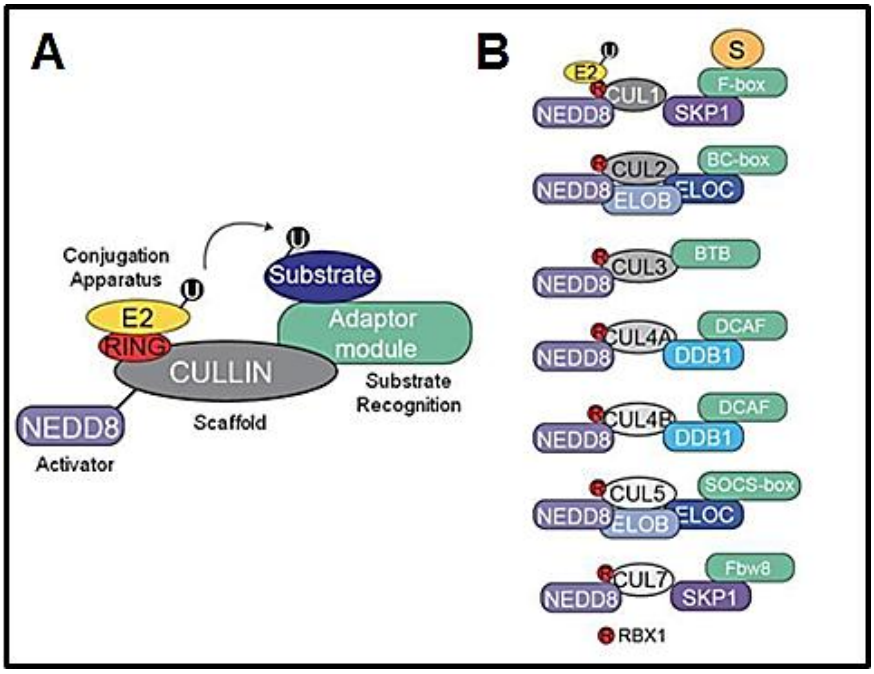


Figure 1.5: The architecture of mammalian CRLs. (A) The basic organization of all CRLs, (B) The composition of the 7 different types of multi-subunit CRLs is illustrated. The substrate recognition subunits unique to each cullin protein are shown in green. Image was adapted from Eric Bennett's faculty webpage UCSD http://biology.ucsd.edu/faculty/bennett.html.

Cullin proteins are also subject to post-translational modification themselves by the small ubiquitin-like modifier Neural precursor cell-expressed developmentally down-regulated protein 8/Related to Ubiquitin1 (Nedd8/RUB1) as shown in Figure 1.6. Nedd8 is conjugated to Cullin proteins by a series of enzymatic reactions involving Nedd8-Activating Enzyme (NAE1) along with a Uba3 subunit, an E2 ubiquitin-conjugating enzyme and an E3-ligase (Rbx1 or DCN1). (Hori *et al.*, 1999; Gong *et al.*, 1999; Gray *et al.*, 2002; Kurz *et al.*, 2005; Hochstrasser *et al.*, 2009).

The conjugation of Nedd8 to Cullin proteins is essential for the function of their respective E3 Cullin RING-ligases (CRLs) ability to target substrate proteins for degradation (Hotton *et al.*, 2008).

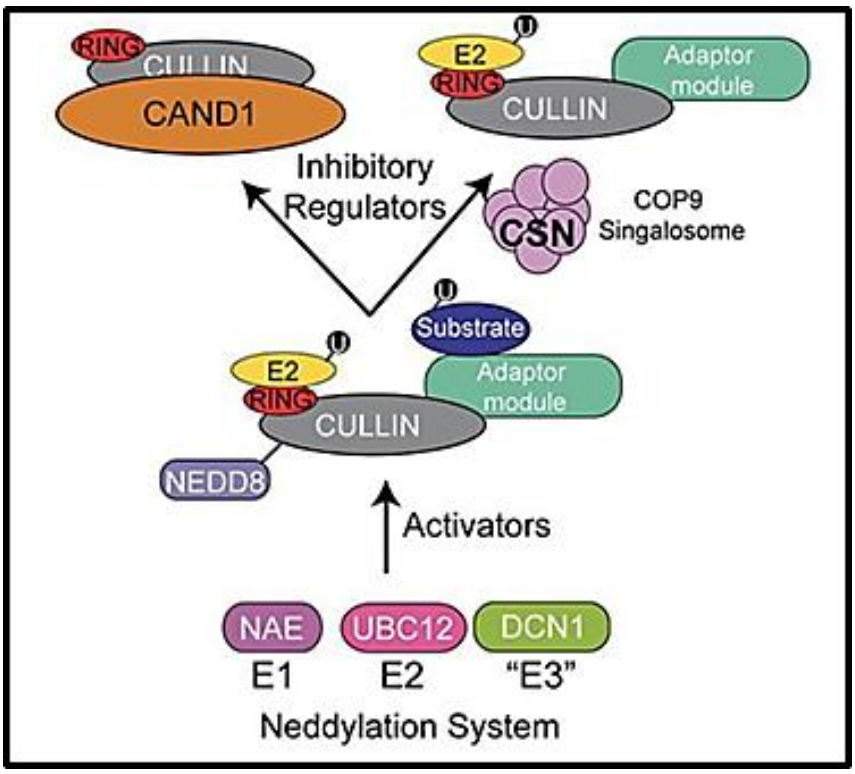


Figure 1.6: Regulation of CRLs by post-translation modification. Illustration of the inhibition of CRL activity by either the prevention of Neddylation by the binding of CAND1 to a cullin protein or the removal of Nedd8 by the COP9 Signalosome. The activation of CRLs by the enzymes responsible for the Neddylation of cullin proteins is termed the 'Neddylation System.' Image was adapted from Eric Bennett's faculty webpage UCSD http://biology.ucsd.edu/faculty/bennett.html.

Inhibiting the neddylation of the Cullin subunits of CRLs would prevent the turn-over of their target proteins and provide a means of achieving CRL-regulated protein homeostasis in cells. The following three sections will introduce the use of small molecules in order to modulate various pathways influencing the cell fate decisions made by stem cells.

1.6 Chemical compounds and small molecules for controlling cell fate

Stem cell self-renewal provides a continuous supply of newly differentiated cells that can re-populate various tissues and organs throughout early development and in adulthood (Vickaryous and Hall, 2006). This process is essential for both development and the regeneration of tissues throughout an organism's lifetime. The

ability of pluripotent stem cells to form any of the three germ layers (mesoderm, endoderm and ectoderm) has raised a considerable amount of interest in the field of regenerative medicine in recent years. In particular, the field of chemical biology has focused much of its efforts on the pursuit of chemical compounds and small molecules that provide a non-genetic means of controlling stem cell fate decisions (Ding *et al.*, 2003; Ding and Schultz, 2004; Schugar *et al.*, 2008; Li and Ding, 2010; Peltier *et al.*, 2010; Lyssiotis *et al.*, 2011; Li *et al.*, 2012).

In order to identify molecules capable of controlling cell fate decisions, it is necessary to understand both the process of cell fate choice as it is made during development and the underlying biochemical mechanisms responsible for these decisions. This section is subdivided into four sections in which I will describe the various cell fate decisions that a stem cell makes throughout development (1.6.1), the use of small molecules to prevent stem cell differentiation (1.6.2), the molecules that have been identified as being able to direct differentiation toward specific lineages (1.6.3) and molecules that demonstrate the ability to increase the reprogramming of somatic cells to a pluripotent state (1.6.4).

1.6.1 Cell fate decisions during fetal development

A tight control over cell fate decisions is essential for proper development of the embryo, tissue regeneration in the adult and the prevention of diseases such as cancer. In humans, all of the various cell types of the body arise from a single totipotent zygote (Vickaryous and Hall, 2006).

The specification of these cells into the tissue-specific cell types that comprise the entire organism occurs in a co-ordinated series of cell fate choices that are determined by both space and time (Hogan, 1999; Gurdon and Bourillot, 2001).

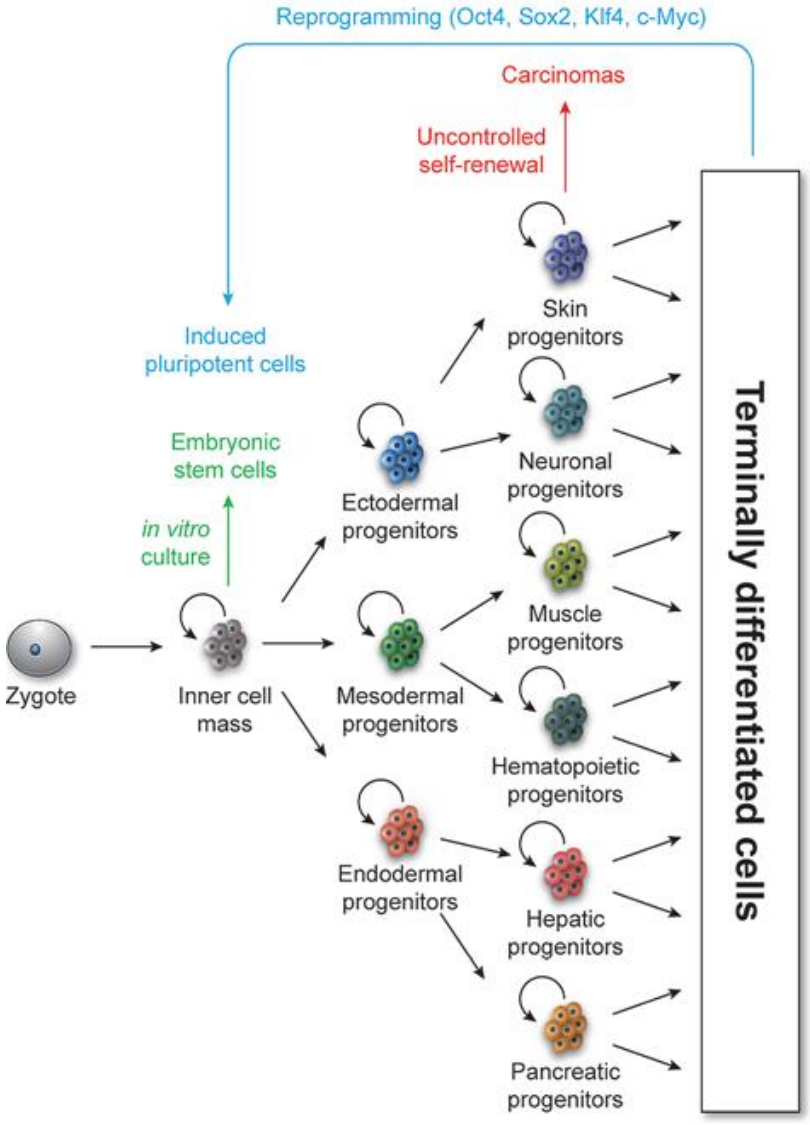


Figure 1.7: Schematic representation of cell lineage commitment during development. The cell fate decisions made by stem cells at various stages of embryogenesis are illustrated. Green text highlights the stage of development represented by current ES model cell lines. Red text highlights the consequence of improper cell specification. Blue text highlights that the entire process of development can be reversed.

Many of the signaling pathways that govern cell type specification and tissue patterning in the mammalian embryo have been identified and are selectively activated in order to control proper fetal development as well as post-natal tissue

homeostasis (Burdon *et al.*, 2002; Bolos *et al.*, 2007; Jiang *et al.*, 2008; Wu and Hill, 2009).

Chemical compounds and small molecules can be employed as molecular probes in order to further dissect the signaling pathways that control cell fate decisions. In this way, they can identify new molecular targets involved in the timing and specification of stem cells both in culture and *in vivo*.

1.6.2 Chemical compounds and small molecules that sustain self-renewal

The long term maintenance of pluripotent stem cells in culture is a challenge for stem cell biologists and has resulted in the pursuit of reagents that promote their self-renewal. Since a shortened G1-phase is characteristic of ES cells, it is generally assumed that this is necessary in order to avoid exposure to mitogenic stimuli that promote differentiation (Burdon *et al.*, 2002; White and Dalton, 2005; Ballabeni *et al.*, 2011). In standard culture conditions of mouse embryonic stem cells, protein factors that promote a rapid G1-phase transit during the cell cycle progression are used to maintain pluripotency. Namely, these two factors are: LIF (Leukemia Inhibitory Factor), a cytokine that activates the STAT3 (Signal Tranducer and Activator of Transcription 3) self-renewal pathway through its binding to the gp130 receptor as well as the MAPK pro-differentiation pathway, and BMP4 (Bone Morphogenic Protein 4) which attenuates the MAPK pathway by inducing the expression of the Id (Inhibitor of differentiation) proteins (Smith *et al.*, 1988; Williams *et al.*, 1988; Niwa *et al.*, 1998; Matsuda *et al.*, 1999; Ying *et al.*, 2003). While the use of exogenous factors and serum were generally well-suited for research applications, the batch-to-batch variations in LIF activity and serum growth

factors yielded inconsistent cellular responses (Ying *et al.*, 2008). Additionally, the use of fetal bovine serum in stem cell culture prevented their use in cell-based therapeutics. Therefore, an animal product-free method of culturing stem cells was required by the field of regenerative medicine.

Using a high-throughput chemical screen of heterocyclic compounds on ES cells and monitoring their expression of Nanog, Oct4, Sox2 and alkaline phosphatase, the first small molecule shown to promote self-renewal of stem cells was discovered. That is, Pluripotin, a dual inhibitor of both RasGAP and ERK1 kinases (Chen *et al.*, 2006). This molecule demonstrated for the first time that stem cells cultured in the presence of a chemical compound rather than protein factors could contribute to all three germ layers in chimeric mice.

In 2008, the fields of chemical biology and stem cell biology were merged once again in an attempt to recapitulate the ability of the BMP4 protein to maintain self-renewal by inhibiting differentiation (Ying *et al.*, 2003). Small molecules inhibiting MEK1 (PD184352) and ERK1/2 (SU540) within the pro-differentiation MAPK pathway were used along with the inhibition of GSK-3β (CHIR99021) in order to release cyclins from suppression (Ying *et al.*, 2008). Using this cocktail of chemical inhibitors, it became possible to cultivate stem cells in absence of serum and extrinsic factors for the first time. As a result, stem cells cultured in the presence of these inhibitors were more amenable to cell-based regenerative therapies and yielded more reproducible and reliable experimental results.

1.6.3 Chemical compounds and small molecules that induce differentiation

In contrast to molecules that sustain pluripotency, compounds have been identified that induce differentiation of stem cells into specialized tissue-specific progenitor cells. A high-throughput chemical screen for compounds that induce ES cell differentiation into definitive endoderm identified IDE1 and IDE (Borowiak *et al.*, 2009). Although the exact mechanism(s) by which these compounds act is unknown, they were found to be more efficient than the commonly used protein factors Activin A and Nodal at driving ES cell differentiation into endoderm lineages (Borowiak *et al.*, 2009). An additional chemical screen was subsequently performed in order to identify compounds that could further direct these chemically-induced endoderm cells to commit into the pancreatic lineage (Chen *et al.*, 2009). This study identified that the sequential treatment of ES cells with IDE1/IDE2 followed by indolactam V was able to induce the differentiation of both mouse and human ES cells into pancreatic progenitor cells. These chemically induced pancreatic stem cells were found to be capable of repopulating the developing gut and give rise to insulin-secreting pancreatic cells following their injection into mouse embryos.

Several chemical library screens have been performed in recent years and have identified novel compounds capable of inducing differentiation of ES cells along various lineages. For example, cardiogenic compounds that generate progenitor cells capable of repairing myocardial repair in rats (Sadek *et al.*, 2008) and the identification of neuropathiazol capable of directing neural progenitor cells into functional neurons (Warashina *et al.*, 2006), to name a few.

1.6.4 Chemical compounds and small molecules that facilitate reprogramming

The demonstrated ability of terminally differentiated mouse and human somatic cells to be reprogrammed by the ectopic expression of four transcription factors provided the possibility of generating patient-specific stem cells for regenerative medicine therapies (Takahashi *et al.*, 2006, Takahashi *et al.*, 2007; Yu *et al.*, 2007). However, fully realizing this potential required further development due to the highly inefficient method of retrovirus-mediated reprogramming. In the initial studies performed by Takahashi *et al.*, 2007 and Yu *et al.*, 2007, only 0.01% of transduced somatic cells generated induced Pluripotent Stem Cells (iPSCs) following several weeks of repeated infections. Other limitations included the integration of viral vectors into the host cell genome as well as the risk of the retroviral promoters driving the reprogramming factors being re-activated and causing cancer in the patient as observed in the mouse (Okita *et al.*, 2007).

In recent years, many of these risks have been eliminated by the use of non-viral vectors (Okita *et al.*, 2008) and without any vector integration into the genome (Kaji *et al.*, 2009; Stadtfeld *et al.*, 2008) and increased reprogramming efficiency using chemical compounds and small molecules (Huangfu *et al.*, 2008; Mikkelsen *et al.*, 2008; Maherali and Hochedlinger, 2009; Li *et al.*, 2009). In the study performed by Huangfu *et al.*, 2008 the use of Valproic acid, an HDAC inhibitor increased reprogramming efficiency using the four factors 400-fold. Furthermore, this compound allowed for iPSCs to be generated using only Oct4 and Sox2. A 100-fold increase in iPSC colony formation has been reported following an appropriately timed treatment with 5-azacytidine, a DNA methyltransferase inhibitor (Mikkelsen *et al.*, 2008). Other molecules involved in the reversal of epigenetic modifications and

gene silencing have also been identified as facilitating the reprogramming process. They are, the histone methyltransferase inhibitors BIX-01294 and Parnate (Shi *et al.*, 2008; Li *et al.*, 2009). The use of BIX-01294 was found to reverse the epigenetic silencing of the Oct4 gene and relieve the requirement for exogenous Oct4 during the reprogramming process (Feldman *et al.*, 2006; Kubicek *et al.*, 2007).

In mouse somatic cells, inhibiting the TGF-β pathway using SB-431542 has been shown to increase reprogramming efficiency 30-fold and greatly reduces the time required for iPSC generation from weeks to days (Maherali and Hochedlinger, 2009; Ichida *et al.*, 2009). Human somatic cells, however, require the combined treatment using both the TGF-β and MEK inhibitors (SB-431542 and PD0325901) to yield a 100-fold increase in iPSC production (Lin *et al.*, 2009).

While several advances have been made in increasing the efficiency of reprogramming somatic cells, and several molecules capable of replacing the requirement of certain factors have been identified, a chemical cocktail capable of generating iPSCs alone has yet to be developed. Further investigation of small molecules and their ability to further enhance reprogramming efficiency as well as replace the various factors required for this process will continue to provide novel methods of generating iPSCs free from the marks left on genomic DNA following genetic manipulation.

1.7 Thesis outline and structure

Presented in this study will be two data chapters each of which having its own separate research focus.

In Chapter 3: Pluripotency Factor Protein-Protein Interaction Mapping, an IP-LC-MS/MS strategy using whole cell extracts as the starting material will be employed in order to further characterize the protein interactomes of several pluripotency factors. While the protein interactomes for several of the chosen factors have been previously identified, those studies provided a snapshot of the interactions occurring solely within the nucleus of ES cells. Additionally, the Pardo *et al.*, 2010 and Cheong *et al.*, 2011 studies identified CUL4B as an interactor with the protein interaction landscape in both the cytoplasm and the nucleus, respectively. Therefore, the proteomic strategy employed in Chapter 3 of this study aims specifically to identify proteins involved in the regulation of pluripotency factors by post-translational modifications.

The overall goal of the research presented in Chapter 3 will be to ascertain pluripotency factor protein-protein interaction networks that extend beyond the nucleus in order give a new perspective on how the highly connected proteomic landscape governing "stemness" is regulated by post-translational modifications in mouse embryonic stem cells.

In Chapter 4: Exploring the Post-Translational Modification of Nanog and Sox2 Using Chemical Compounds and Small Molecules, both the stability and the interactions of the Nanog and Sox2 proteins will be examined following targeted inhibition of the kinases and cullin proteins identified in Chapter 3 of this study. While the dimerization, phosphorylation and stabilization of Nanog has been identified by Mullin *et al.*, 2008, Wang *et al.*, 2008 and Moretto-Zita *et al.*, 2010, the kinase responsible for this phosphorylation has not been identified. Therefore, one of the aims of Chapter 4 in this study will be to inhibit all of the kinases identified in

Chapter 3 of this study in order to reveal the kinase or kinases responsible for phosphorylating Nanog. A compound found to inhibit the phosphorylation and stabilization of Nanog will provide a non-genetic means to homogeneously prime an entire population of ES cells for subsequent differentiation into any cell type desired. At the same time, this compound would serve as a potent anti-cancer therapeutic since Nanog is frequently re-expressed in the most aggressive forms of cancer (Jeter *et al.*, 2009; Zbinden *et al.*, 2010; Brandner *et al*., 2010; Meng *et al.*, 2010).

Conversely, chemical compounds found to elevate phosphorylated Nanog would provide a novel method of stabilizing the specific differentiation-resistant phenotype that Nanog confers to ES cells as well as to somatic cells being reprogrammed (Chambers *et al.*, 2007; Yu *et al.*, 2007). Specifically, cullin proteins identified as part of the pluripotency interaction networks obtained in Chapter 3 of this study will be inhibited both chemically using MLN4924 and individually using shRNA. The aim of this series of experiments will be to provide insight into whether or not Nanog and Sox2 protein levels are regulated by the UPS in embryonic stem cells. Further to this end, the inhibition of specific cullin proteins will help to identify the specific cullin-RING E3 ligase responsible for targeting these master pluripotency factors. The identification of these CRLs will allow for the development of targeted inhibitors specifically inhibiting either the degradation of Nanog or Sox2. These reagents will allow for the reversible and dosage-dependent activation of the core pluripotency transcriptional circuitry without leaving the footprints of genetic manipulation upon the cells.

The overall goal of the research presented in Chapter 4 of this study is to identify novel chemical compounds and small molecules that allow for the

modulation of the Nanog and Sox2 proteins thus, providing the fields of regenerative medicine and oncology with the first specific and targeted reagents controlling both the maintenance of pluripotency and the induction of differentiation in cells.

CHAPTER 2: Materials and Methods

2.1 Materials:

2.1.1 Kits and reagents:

2.1.1.1 Kits

Spin Miniprep kit	QIAGEN
Endo-Free Maxi kit	QIAGEN
Protein A/G plus agarose conjugation kit	Pierce
NativePAGE™ sample prep kit	Invitrogen

2.1.1.2 Reagents

NuPAGE® Novex® Bis-Tris 4-12% gradient gel	Invitrogen
NuPAGE® LDS sample loading buffer, 4X	Invitrogen
SeeBlue® pre-stained protein standard	Invitrogen
BenchMark pre-stained protein marker	NEB
SimplyBlue™ Safe Stain (Coomassie)	Invitrogen
NativePAGE™ (3-12%) gradient gel	Invitrogen
NativePAGE™ 4X sample loading buffer	Invitrogen
NativePAGE™ 20X cathode buffer	Invitrogen
NativePAGE™ cathode buffer additive	Invitrogen
NuPAGE® Transfer buffer	Invitrogen
NativePAGE™ 5% G-250 Sample Additive	Invitrogen
NativeMark™ unstained protein standard	Invitrogen
anti-FLAG® M2 agarose	Sigma
3XFLAG® peptide	Sigma
LIF (Leukemia Inhibitory Factor), >95%	Sigma
DMA (N,N-Dimethylacrylamide), 99%	Sigma
anti-HA High Affinity Matrix	Roche
Protein A/G Plus agarose	Pierce
NuPAGE® MOPS running buffer	Invitrogen
NuPAGE® MOPS Transfer buffer	Invitrogen

SuperSignal™ West Pico/Femto Chemiluminescent Substrate	Pierce
Amicon® Ultra-0.5mL, centrifugal spin columns for protein concentration and purification, (YM-3).	Millipore
GMEM (Glasgow Minimal Essential Medium)	Sigma
HyClone Fetal Bovine Serum, defined, 40 nm-filtered	Thermo
Puromycin	InvivoGen
DAPI (4', 6-Diamidino-2-phenylindole dihydrochloride)	Sigma
ESGRO COMPLETE™ Gelatin solution, Ultrapure water with 0.1% Gelatin	Millipore
Sepharose CL-4B	Sigma
NP40 (Igepal CA-630)	Sigma
Glycerol	Merck
DTT (Dithiothreitol)	Melford Labs Ltd.
Benzonase Nuclease HC, purity >90%	Novagen
Lipofectamine™ 2000 Reagent, 1mg/mL	Invitrogen
OptiMEM serum-free medium	Invitrogen
Complete, EDTA-free Protease inhibitor Cocktail tablets	Roche
Phosphatase Inhibitor Cocktail Set II, 100X	Calbiochem
0.2 µm vacuum filtration unit	Nalgene
Lab-Tek® II Chamber Slide™ 8 well Glass Slide	ThermoFisher
Cover Glass No.1, 24mm x 60mm	Corning
Paraformaldehyde, 16% (w/v) aq. soln., methanol-free	Alfa Aesar
Triton® X-100 (Polyethylene glycol *tert*-octylphenyl ether)	Fluka
Bovine Serum Albumin (BSA) (Fraction V)	Fisher
ProLong® Gold antifade reagent	Invitrogen
10% Normal Goat Serum, Ready-to-use	Invitrogen
Streptavidin, Alexa Fluor® 555 conjugate	Invitrogen
10X Phosphate –buffered saline (10X PBS)	Invitrogen
λ-phosphatase	NEB

2.1.2 Solutions

PBS:

137mM NaCl, 2.7mM KCL, 4.3 mM Na_2HPO_4, 1.47mM KH_2PO_4

PBS-T:

137mM NaCl, 2.7mM KCL, 4.3 mM Na_2HPO_4, 1.47mM KH_2PO_4, 0.1% (v/v) Tween20

TBS:

50 mM Tris base, 150 mM NaCl. pH adjusted to 7.6

TBS-T:

50 mM Tris base, 150 mM NaCl. pH adjusted to 7.6, 0.1% (v/v) Tween20

IP Lysis buffer:

50 mM Tris-HCL pH 7.5, 250 mM NaCl, 5 mM $MgCl_2$, 0.1% (v/v) NP40,10% (v/v) Glycerol, 1 mM DTT, Benzonase nuclease HC 1µL/mL of lysis buffer, 1X Complete Protease, Inhibitor Cocktail, 1X Phosphatase Inhibitor Cocktail Set III.

IP wash buffer 1:

50 mM Tris-HCL pH 7.5, 250 mM NaCl, 5 mM $MgCl_2$, 0.01% (v/v) NP40,10% (v/v) Glycerol, 1 mM DTT, Benzonase nuclease HC 1µL/mL lysis buffer, 1X Complete Protease, Inhibitor Cocktail, 1X Phosphatase Inhibitor Cocktail Set III.

IP Wash Buffer 2:

50 mM Tris-HCL pH 7.5, 250 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, Benzonase nuclease HC 1µL/mL lysis buffer, 1X Complete Protease, Inhibitor Cocktail, 1X Phosphatase Inhibitor Cocktail Set III.

3XFLAG 5mg/mL Peptide Stock Solution:

Entire contents of 1 vial containing 5mg of 3XFLAG peptide were resuspended in 1mL of buffer containing 50mM Tris-HCl, pH 7.4 and 150mM NaCl. Solution aliquotted and stored at -20°C.

3XFLAG 200µg/mL peptide elution solution:

5mg/mL 3XFLAG stock solution diluted to a final working concentration of 200µg/mL in Wash Buffer 2.

Western blot blocking buffer:

5% (w/v) skimmed milk powder dissolved in TBS-T or PBS-T.

Western blot antibody buffer:

1% (w/v) skimmed milk powder dissolved in TBS-T or PBS-T.

Western blot washing solution:

1X TBS or 1X PBS solution containing 0.1% (v/v) Tween20.

Native Anode buffer solution:

NativePAGE™ Running buffer, 20X was diluted to 1X in ddH_2O.

Native Dark Blue Cathode buffer solution:

10 mL of NativePAGE™ [20X] Cathode additive was added to 900 mL of 1X NativePAGE™ Running buffer.

Native Light Blue Cathode buffer solution:

1 mL of NativePAGE™ [20X] Cathode additive was added to 999 mL of 1X NativePAGE™ Running buffer.

NativePAGE™ reducing solution:

4X NuPAGE® LDS sample loading buffer was diluted to 1X with ddH_2O and 50 mM DTT was added.

NativePAGE™ alkylating solution:

4X NuPAGE® LDS sample loading buffer was diluted to 1X with ddH_2O and 50 mM DMA was added.

NativePAGE™ quenching solution:

4X NuPAGE® LDS sample loading buffer diluted to 1X with ddH_2O and 5 mM DTT, 20% (v/v) ethanol were added.

IF (Immunofluorescence) fixing solution:

4% (w/v) paraformaldehyde solution in 1X PBS.

IF (Immunofluorescence) cell permeabilization solution:

0.5% (v/v) Triton® X-100 in 1X PBS.

IF (Immunofluorescence) blocking solution:

10% (v/v) Normal Goat Serum, Ready-to-use (Invitrogen).

10,000X DAPI long term stock solution in ddH_2O:

10 mg of DAPI was dissolved in 1 mL of ddH_2O and stored at -20°C.

1000X DAPI working solution in ddH_2O:

A (1:10) dilution in ddH_2O was prepared from the 10, 000X DAPI stock solution and stored at -20°C.

IF (Immunofluorescence) DAPI staining solution:

Immediately prior to use, 1000X DAPI was diluted to 1X in PBS.

Ponceau S staining solution:

0.2% (w/v) Ponceau S, 3% (w/v) trichloroacetic acid, 3% (w/v) sulfosalicylic acid.

50 mM ammonium bicarbonate solution, (NH_4HCO_3):

197mg NH_4HCO_3 in 50mL MilliQ water; 0.2 µm-filtered

200mM DTT:

1.54mg per 50µL of water.

500mM Iodoacetamide**:**

1 mg per 10µL of water.

Trypsin digestion solution:

20 µg of sequencing grade trypsin (Promega) dissolved in 1 mL of

50 mM ammonium bicarbonate solution.

20 mM DTT solution:

200 mM DTT diluted (1:10) in 50 mM ammonium bicarbonate solution.

50 mM iodoacetamide solution:

500 mM Iodoacetamide diluted (1:10) 50 mM ammonium bicarbonate solution.

2.1.3 Plasmids

pPyCAG-IP/3XFLAG-Nanog	(a gift from Tilo Kunath)
pCAG-6XHA-Sox2-IP	Maruyama *et. al.*, 2005
pPyCAG-6XHA-Pik3cd-IP	Takahashi *et. al.*, 2003

2.1.4 Cell lines

Cell Line	Description	Reference/Source
E14tg2a	Wild type, hypoxanthine phosphoribosyltransferase deficient, mouse embryonic stem cell line.129/Ola background. (Clonal)	Hooper *et al.*, 1987
TNG	E14Tg2a mouse embryonic stem cell line in which the endogenous Nanog gene has been replaced by the eGFP gene. eGFP was inserted between the homology arms precisely at the Nanog AUG codon. eGFP expression is linked through an IRES to puromycin resistance encoded by the pac gene.	Chambers *et al.*, 2007
3XFLAG-Nanog	E14tg2a stably expressing 3XFLAG-Nanog from the pPyCAG-IP plasmid. 3XFLAG-Nanog expression is linked through an IRES to puromycin resistance encoded by the pac gene. (Clonal)	This study
6XHA-Sox2	E14tg2a stably expressing 6XHA-Sox2 from the pCAG-IP plasmid. 6XHA-Sox2 expression is linked through an IRES to puromycin resistance encoded by the pac gene. (Clonal)	This study
6XHA-Pik3cd	E14tg2a stably expressing 6XHA-Pik3cd from the pCAG-IP plasmid. 6XHA-Pik3cd expression is linked through an IRES to puromycin resistance encoded by the pac gene. (Clonal)	This study

2.1.5 Lentiviral particles for shRNA

Dharmacon SMARTvector 2.0 lentiviral shRNA vial particles from a pool of three different gene targeting sequences for the same gene to ensure successful knockdown of the targeted gene. Pre-packaged CUL3, CUL4B and CUL7-targeting lentiviral shRNA particles were purchased from Thermo Scientific/Dharmacon (distributed by Abgene Ltd., UK). Table 2.1 lists the source clone IDs, as well as, the targeting sequences of each for each construct in the pre-packaged shRNA pools used.

Catalogue No.:	Source Clone ID:	Gene Symbol:	Gene Target Sequence:
VSM5514	SH-062951-01-10	CUL3	GAGATCAAGTTGTACGGTA
VSM5514	SH-062951-02-10	CUL3	GATTGCCAGAGAGCGGAAA
VSM5514	SH-062951-03-10	CUL3	TTCAAGAAATCCAGCGTAA
VSM5514	SH-041281-01-10	CUL4B	GGTGAACACTTAACGGCAA
VSM5514	SH-041281-02-10	CUL4B	AACTAATCGGCTTTACGCA
VSM5514	SH-041281-03-10	CUL4B	GGGCTCATATTATAAGTGA
VSM5514	SH-054741-01-10	CUL7	GGGTGCTCAAGATTCGAGA
VSM5514	SH-054741-02-10	CUL7	AATTGCCTTGTTGTCCGAA
VSM5514	SH-054741-03-10	CUL7	ATGTTGACCGGCTCGTCTA

Table 2.1 List of pre-packaged lentiviral shRNA viral particles used in this study for knockdown of cullins in mESCs.

2.1.6 Antibodies

Primary Antibodies:

Mouse monoclonal anti-FLAG® M2 antibody	Sigma
Rat monoclonal anti-HA High Affinity antibody	Roche
Rabbit anti-Nanog antibody	Bethyl
Monoclonal anti-human/mouse Sox2 antibody	NEB
Anti-Sox2-Alexa® 555 conjugate antibody	R&D Systems
Monoclonal anti-human ERRβ/NR3B2 antibody	Perseus Proteomics
Goat anti-SALL4 antibody	LifeSpan
Rabbit monoclonal against CUL4B (clone EPR3200)	Origene
Rabbit anti-human CUL4B monoclonal (EPR3200)	LifeSpan
Rabbit anti-CUL7 antibody	Bethyl
Rabbit polyclonal anti-pHH3 antibody	NEB

Secondary Antibodies:

Anti-mouse-HRP antibody	GE Amersham
Anti-rat-HRP antibody	GE Amersham
Anti-rabbit-HRP antibody	GE Amersham
Anti-sheep/goat-HRP antibody	GE Amersham
Anti-mouse-Alexa® 488 antibody	Invitrogen
Anti-rabbit- Alexa® 488 antibody	Invitrogen
Anti-mouse- Alexa® 594 antibody	Invitrogen
Anti-rabbit- Alexa® 594 antibody	Invitrogen
Anti-mouse- Alexa® 647 antibody	Invitrogen
Anti-rabbit- Alexa® 647 antibody	Invitrogen

2.1.7 Chemical compounds and small molecules

Unless otherwise stated, all compounds were purchased from Calbiochem/Merck.

Compound:	Final concentration in DMSO:
Nocodazole	[20 µM]
Br-cAMP	[50 nM]
PKI (14-22)	[50 µM]
Ht-31	[50 µM]
MLN-4924	[1 µM]
H89	[10 µM]
Forskolin	[10 µM]
Wortmannin	[10 µM]
Aurk*i*III	[10 µM]
3*i* cocktail	1X from 100X stock solution (Stem Cell Sciences)

2.1.8 Culture medium:

2.1.8.1 Bacterial culture medium

Luria-Bertani Medium (LB) (Miller, 1972) Difco Bacto tryptone (10 g/L), Difco Bacto yeast extract (5 g/L), NaCl (5 g/L). Prior to autoclaving the pH was adjusted to 7.2. Solid LB agar plates were made by adding 2% (w/v) Difco Bacto agar to the media.

2.1.8.2 Cell culture medium

Regular Growth Medium (for 500mL bottle):

500 mL	GMEM	Sigma
50 mL	HyClone FBS, ES cell defined, 40 nm-filtered	Thermo
5.5 mL	100X L-Glutamine solution	GIBCO
5.5 mL	100X MEM NEAAs solution	GIBCO
5.5 mL	100X Sodium pyruvate solution	GIBCO
1.1 mL	[50mM] 2-Mercaptoethanol solution	GIBCO
567.6 µL	1000X LIF, >95%	Sigma

2.2 Methods:

2.2.1 DNA techniques:

2.2.1.1 Plasmid transformation into DH5α

50 – 100 ng of plasmid DNA was mixed with 50uL of chemically competent DH5α *E. coli* cells. This mixture was incubated on ice for 30 minutes, then at 42°C for 45 seconds and returned to ice for 2 minutes. 500 µL of LB medium containing 2% (w/v) glucose was added to the transformation mixture and allowed to incubate at 37°C with agitation at 225 rpm for 1 hour. Cells were then plated onto LB agar plates

supplemented with the appropriate selection antibiotic and allowed to grow O/N at 37°C. Ampicillin was used at a final working concentration of 100 µg/mL.

2.2.1.2 Plasmid isolation from DH5α

A single colony of *E. coli* transformed with plasmid DNA (Section 2.2.1.1 Plasmid transformation into DH5α) was picked and inoculated to 5 mL of LB medium containing the appropriate selection antibiotic. This culture was allowed to grow O/N at 37°C with agitation at 225 rpm. The following day, bacterial pellets were harvested by centrifugation at 10, 000 x g for 2 minutes at 4°C. Plasmid DNA was isolated from bacterial cell pellets using the QIAGEN Spin Miniprep kit reagents and following the protocol detailed in the QIAprep Miniprep Handbook.

2.2.1.3 Large scale plasmid preparation for transfection into mESCs

A single colony of *E. coli* transformed with plasmid DNA (Section 2.2.1.1 Plasmid transformation into DH5α) was picked and inoculated to 5 mL of LB medium containing the appropriate selection antibiotic. This culture was allowed to grow O/N at 37°C with agitation at 225 rpm. The following day, 300 µL of the O/N bacterial culture was inoculated into a 1L Erlenmeyer flask containing 300 mL of LB medium containing the appropriate selection antibiotic. Cultures were grown at 37°C with agitation at 225 rpm until an OD of 0.8 was reached. Bacterial pellets were harvested by centrifugation at 6, 000 x g for 10 minutes at 4°C. Plasmid DNA was isolated from bacterial cell pellets using the QIAGEN Edo-Free Maxi kit reagents and following the protocol detailed in the QIAGEN plasmid purification handbook.

2.2.2 Culture of mouse embryonic stem cells (mESCs)

2.2.2.1 Routine culture and passaging of mESCs

Unless otherwise stated, mESCs were maintained in as feeder-independent monolayer cultures on gelatin coated tissue culture flasks as per Williams *et al.*, 1988 and Smith *et al.*, 1988.

Cell harvesting and passing protocol:

Medium was removed from cells cultured in a T25 flask. Medium was rinsed from the surface of monolayer by adding 5 mL of 1X D-PBS (GIBCO). PBS was aspirated from the flask and 2 mL of 0.05% (w/v) trypsin-EDTA, 1X solution (GIBCO) was added. Cells were incubated with 1X trypsin-EDTA solution at RT for 5 minutes. Cells were dislodged by gently tapping the side of the flask. Trypsin was neutralized by adding 2 mL of regular growth medium. Cells were re-suspended into a single-cell suspension by pipetting up and down several times with a 5 mL pipette. The cell suspension was transferred to a 50 mL centrifuge tube already containing 10 mL of pre-warmed regular growth medium. Cells were pelleted from suspension by centrifugation at 300 x g for 5 minutes at 4°C. The supernatant was aspirated from pelleted cells which were then re-suspended cells in 2 mL of growth medium. The appropriate volume (see **Note 1)** of cell suspension was dispensed into a T-25 flask coated with gelatin using ESGRO COMPLETE™ Gelatin solution (Millipore) containing 10 mL of growth medium. In general, cells were passed when the culture reached 80% confluency and the medium and growth factors were replaced every 2 days.

1) For routine culture, cells were plated at a [1:4] dilution into the appropriate medium. For over the weekend cultures, cells were plated at a (1:20) dilution into appropriate medium.

2.2.2.2 Freezing of mESCs

Cells grown to 80% confluency in T25 flasks were harvested as per protocol detailed in Section 2.2.2.1 Routine culture and passaging of mESCs and re-suspended in 2 mL of mESC freezing medium. One mL of mESCs in freezing medium was added to a 1.5 mL Cryotube (Corning) and immediately placed on dry ice. mESCs were transferred to -80°C for storage.

mESC Freezing Medium (for 10 mL):

9mL	Regular growth medium	
1 mL	DMSO	Sigma

2.2.2.3 Transfection of mESCs

The day before transfection, cells were harvested and plated into gelatin-coated 6-well plates at a dilution into fresh regular growth medium to provide a flat monolayer of mESCs of 40-50% confluency on the day of transfection. On the day of transfection, growth medium was removed and 2 mL of fresh growth medium per well of a 6-well plates was added to cells. One µg of plasmid DNA diluted in RT OptiMEM™ serum-free medium (GIBCO) was complexed into liposomes for transfection using Lipofectamine™ 2000 reagent (Invitrogen) according to manufacturer's instructions. After 4 hours of incubation at 37°C and 7.5% CO_2, 2mL of regular growth medium was added to each well of the 6-well plate containing transfection complexes. Cells were returned to 37°C and 7.5% CO_2 and allowed to recover overnight. The following day, cells were harvested and plated at a (1:10) dilution into gelatin-coated 6-well plates containing 2 mL of PURO selection medium containing 1 µg/mL puromycin (InvivoGen). Cells were maintained in

selection medium until the appearance of single colonies occurred. The lines established from the individual colonies were expanded in culture in order to test for expression of the epitope-tagged recombinant proteins.

PURO Selection Medium (for 500mL bottle):

500 mL	GMEM	Sigma
50 mL	HyClone FBS, ES cell defined, 40 nm-filtered	Thermo
5.5 mL	100X L-Glutamine solution	GIBCO
5.5 mL	100X MEM NEAAs solution	GIBCO
5.5 mL	100X Sodium pyruvate solution	GIBCO
1.1 mL	[50mM] 2-Mercaptoethanol solution	GIBCO
567.6 µL	1000X LIF, >95%	Sigma
56.8 µL 1	10 mg/mL Puromycin	InvivoGen

2.2.2.4 Large-scale culture of mESCs for IP-LC-MS/MS studies

Mouse embryonic stem cells were passaged and maintained as described in Section 2.2.2.1 Routine culture and passaging of mESCs. Routine culture and passaging of mESCs with the exception of being scaled up to 4 x T-175 flasks per twenty 20mm x 150mm tissue culture dishes to be plated. That is, ten 20mm x 150mm dishes per IP, duplicated for each epitope-tagged bait protein. Cells were grown to near 100% confluency on gelatin-coated T-175 flasks. Cells from 1 x T-175 flask were harvested as described in Section 2.2.2.1 Routine culture and passaging of mESCs, resuspended in 5mL of regular growth medium and plated (1:5) into five 20mm x 150mm Cell^{+} Tissue Culture Dishes containing 30 mL of growth medium. After culturing for 48 hours at standard culture conditions, the dishes were 80-90% confluent and harvested for subsequent immunopurification experiments.

2.2.2.5 Culture of mESCs in hypoxia

Protocols are the same as those described in Section 2.2.2.1 Routine culture and passaging of mESCs with the following exceptions: All hypoxia work was performed in the XVivo model G300C hypoxia workstation. All incubators attached to the workstation were controlled to 5% O_2 and 7.5% CO_2 in a completely automated fashion using the XVivo O_2 control Watview software. For all hypoxia experiments in this study, the levels of oxygen and carbon dioxide were fixed at 5% and 7.5%, respectively. Additionally, the entire enclosed workspace was maintained under the same hypoxic conditions which ensured that all cultures and materials within the workstation were constantly maintained in the same low oxygen environment. Prior to beginning any hypoxia work, all plastic consumables to be used were placed inside the workstation under hypoxic conditions in order to purge oxygen from all plastics to come in contact with the cells and/or growth medium and washing buffers. All required 500 mL bottles of 1X D-PBS were placed inside the workstation with lids unscrewed and resting atop the bottle. This ensures that all oxygen in the 1X D-PBS used to wash the cells has been sufficiently purged from the solution. All perishable reagents (i.e., regular growth medium, trypsin, etc.) were aliquotted in small volumes (not greater than 50 mL) into wide-mouth containers and placed in the workspace under hypoxic conditions for a minimum of 3 hours at RT in order to equilibrate them to the desired 5% oxygen. Times required for sufficient oxygen displacement were determined and recommended by Matt Freeman, engineer at BioSpherix).

2.2.2.6 Lentiviral transduction of mESCs

Using the pre-made lentiviral particles described in Section 2.1.5 Lentiviral particles for shRNA, the 3XFLAG-Nanog mESC cell line was transduced following the protocol described in Embryonic Stem Cells: methods and protocols by Kursad Turksen, 2002. Briefly, mESCs were plated on gelatin-coated 6-well plates at a density of $5x10^4$ cells/well 24 hours prior to transduction. The following day, medium was removed from cells and 5 x 10^7 TU lentiviral particles were added directly to the cells. Following 1 minute incubation at RT, 100 µL of regular growth medium containing 10µg/mL of Polybrene was added to cells. mESCs were incubated with this mixture for 1 hour at 37°C, 7.5% CO_2. 2 mL of regular growth medium was added to each well of the 6-well plate and cell were returned to standard culture conditions for 24 hours. After 24 hours, the virus-containing medium was replaced with fresh regular growth medium and cells were allowed to incubate for a further 24 hours. The next day, transduced cells were harvested and plated into puromycin selection medium to select for stable integrants.

2.2.2.7 Cell cycle synchronization of mESCs

Protocol followed as per Savatier *et al.*, 1996 with slight modifications. In brief, cells were treated with [20µM] nocodazole overnight. The following morning, mitotic cells were shaken off by hitting the flask several times with the palm of hand. Cells floating in the medium were then collected by decanting the medium into 50 mL Falcon tubes. Mitotic cells were pelleted by centrifugation at 500 x g for 5 minutes at 4°C. The medium was aspirated off the cell pellets which were gently resuspended in pre-warmed regular growth medium. Tubes were filled with regular

growth medium and centrifuged at 500 x g for 5 minutes at 4°C. Medium was discarded and the cells (now free from traces of nocodazole) were gently resuspended in 1 mL of regular growth medium. Cells were plated at a density of 2.0 x 10^5 cells/well into each well of a 6-well plate. Since mitotic cells require 4-5 hours to strongly adhere to the dish, T0 in these experiments represents 4 hours post-release into nocodazole-free medium.

2.2.3 Protein methods:

2.2.3.1 Preparation of samples for SDS-PAGE

Cell pellets were lysed by the addition of an equal volume of cell lysis buffer in µL per mg of cell pellet. Pellets were re-suspended by vortexing and sheared by pipetting up and down with a P1000 pipette tips. The lysates were allowed to stand on ice for 10 minutes and clarified by centrifugation at 12, 000 x g for 5 minutes. 4X NuPAGE® LDS sample loading buffer was added to final concentration of 1X with 100µM DTT to clarified lysates. Samples were boiled for 5 minutes at 100°C and vortexed briefly prior to loading onto gel.

2.2.3.2 SDS-PAGE for western blotting

Protein samples were separated by SDS-PAGE using NuPAGE® 4-12% Bis-Tris Gel, 1.0 x 10 well (Invitrogen). Gels were run at 200V for 1 hour at RT in 1X NuPAGE® MOPS running buffer (Invitrogen) using an XCell SureLock™ Electrophoresis Cell (Invitrogen).Following electrophoresis, gels were removed from their cassettes and incubated in 1X NuPAGE® Transfer buffer for 20 minutes at RT with orbital shaking prior to electrophoretic transfer.

2.2.3.3 Electrophoretic transfer of SDS-PAGE separated proteins

Two pieces of Extra Thick Blot Paper, Protean xi size (Bio-Rad Laboratories, Inc.) was cut to roughly 10 cm x 10cm in size (2 per gel to be transferred) and nitrocellulose membrane, 0.2 µm (Bio-Rad Laboratories, Inc.) was cut to roughly 9 cm x 9 cm in size. Gels, blotting paper and membranes were prepared for semi-dry electrophoretic transfer by soaking in 1X NuPAGE® Transfer buffer for 20 minutes at RT with orbital shaking. Routine electrophoretic transfer was performed at 15V for 30 minutes at RT using a Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad Laboratories, Inc.).

2.2.3.4 Immunoblotting

Following semi-dry transfer of proteins to membranes in Section 2.2.3.3 Electrophoretic transfer of SDS-PAGE separated proteins, membranes were incubated in western blot blocking solution for 2 hours at RT with orbital shaking. Membranes were washed 3 x 5 minutes at RT with orbital shaking in western blot blocking buffer. All primary antibodies solutions were prepared as a (1:1000) dilution in western blot antibody buffer. All primary antibody buffer incubations were carried out overnight at 4°C with orbital shaking. The next day, membranes were washed 3 x 5 minutes at RT with orbital shaking in western blot washing buffer. All secondary antibody solutions were prepared as (1:10, 000) dilutions in western blot antibody buffer. All secondary antibody solution incubations were carried out for 2 hours at RT with orbital shaking. Membranes were washed 3 x 5 minutes in washed 3 x 5 minutes at RT with orbital shaking in western blot washing buffer. Membranes were incubated with 2 mL of SuperSignal™ West Pico/Femto

Chemiluminescent Substrate (Pierce) prepared as per manufacturer's protocol for 1 min at RT with shaking. Excess substrate was drained from membranes before placing them between two clean Niceday transparencies (Office Depot) and securing them into a FUJIFilm BAS Cassette2. Membranes were exposed to Amersham Hyperfilm ECL High Performance Chemiluminescence film (GE Healthcare) and films were subsequently developed using a SRX-101A tabletop film processor (Konica, U.S.A.).

2.2.3.5 NativePAGE™ for western blotting

This technique was performed on samples of immunopurified 3XFLAG-Nanog protein complexes isolated under native conditions, eluted by peptide competition and concentrated. NativePAGE™ gels are the commercially available equivalent to the BN (Blue-Native) PAGE technique developed by Schägger and von Jagow, 1991. In both cases, the G-250 dye serves as the charge-shift molecule so that the molecular weight native proteins and protein complexes may be determined. Eluates of immunopurified 3XFLAG-Nanog and Nanog-containing complexes were prepared for native PAGE by adding 4X NativePAGE™ sample loading buffer (Invitrogen) to a final concentration of 1X in ddH_2O and NativePAGE™ 5% (w/v) G-250 Sample additive (Invitrogen) to $1/4^{th}$ the detergent concentration. The wells of a 10-well NativePAGE™ (3-12%) gradient gel (Invitrogen) were washed 2 times with dark blue cathode buffer prior to filling the inner buffer chamber completely with the dark blue cathode buffer. A total sample volume of 25 µL was loaded into each well and 10 µL of NativeMark™ (Invitrogen). After filling the outer buffer chamber with 200 mL of anode buffer, gels were run for at 150V for 90 minutes at

RT using an XCell SureLock™ Electrophoresis Cell (Invitrogen). Following electrophoresis, NativePAGE™ gels were transferred to membranes following the protocol for electrophoresis detailed in Section 2.2.3.3. Following transfer, proteins were fixed onto the membranes by incubating the membranes in 8% (w/v) acetic solution for 10 minutes at RT with orbital shaking. Membranes were washed several times in 1X TBS or PBS before proceeding to immunoblotting with anti-FLAG antibody as per the protocol detailed in Section 2.2.3.4.

2.2.3.6 Two-Dimensional NativePAGE™ to SDS-PAGE for western blotting

Samples were prepared and electrophoresis performed in the same manner as that outlined in Section 2.2.3.5. Following electrophoresis, gels were cut into vertical strips (from sample loading well to the bottom of the gel). Each strip was then inserted into a 50 mL Falcon tube containing 10 mL of NativePAGE™ reducing solution for 30 minutes at RT. The reducing solution was removed and 10 mL of NativePAGE™ alkylating solution was added to each strip. Strips were allowed to incubate for 15 minutes at RT. This solution was decanted and replaced with 10 mL of NativePAGE™ quenching solution. Gel strips were incubated with quenching solution for a minimum of 30 minutes at RT to a maximum of overnight at 4°C. Following quenching, each gel strip was removed from quenching solution onto a petri dish and cut into 14 2.5 mm width slices from the top of the strip (sample loading well) to the bottom (approximately 25 kDa band). Gel slices were allowed to dry and shrink at RT for 1 hour prior to loading. The wells of a NuPAGE® 4-12% Bis-Tris Gel, 1.0 x 15 well (Invitrogen) were rinsed with 1X NuPAGE® LDS sample loading buffer containing 100 µM DTT. Each of the 14 gel slices were carefully slid

into a well of the 15-well SDS-PAGE gel. Once all 14 slices and molecular weight marker were inserted/loaded onto the SDS-PAGE gel, all wells were overlayed with 1X NuPAGE® LDS sample loading buffer containing 100 µM DTT. Gel electrophoresis, electrophoretic transfer and immunoblotting were performed as per Sections 2.2.3.2 - 2.2.3.4.

2.2.3.7 Immunopurification of epitope-tagged proteins from mESCs for LC-MS/MS

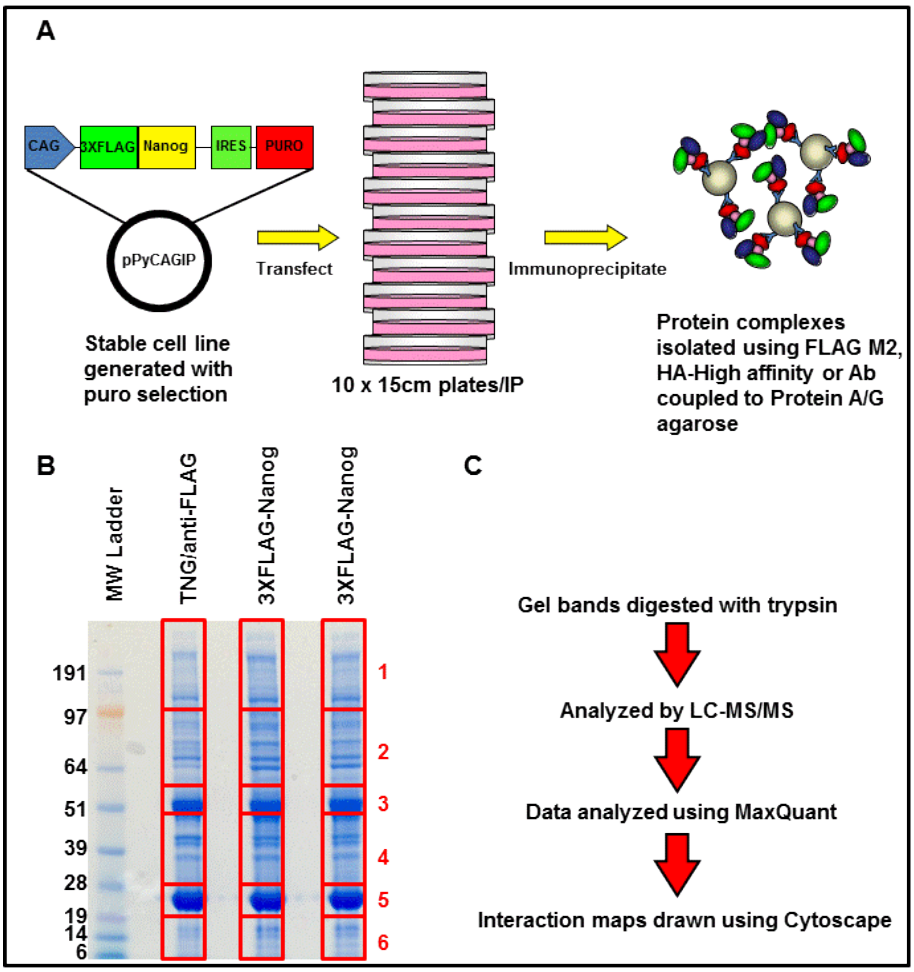


Figure 2.1: The IP-LC-MS/MS-based protein-protein interaction mapping platform established in this study. A) Workflow diagram of recombinant protein expression and immunopurification from mouse ESCs. **B)** Representative Coomassie-stained gel of purified protein complexes. Gel slices excised and digested for LC-MS/MS analysis are indicated by red boxes. **C)** Sample preparation and LC-MS/MS data analysis workflow.

Ten 20 mm x 150 mm Cell^{+} tissue culture dishes (Sarstedt) were harvested per IP of 3XFLAG-Nanog, 6XHA-Sox2 and 6XHA-Pik3cd mESC cells lines and TNG cell line for control IPs. Prior to beginning the harvest, all IP buffers and solutions were filtered using a Nalgene 0.2 µm vacuum filtration unit (NalgNunc Intl.) and placed on ice for a minimum of 30 minutes. Dishes to be harvested were placed on ice, medium aspirated and washed 2 times with 20 mL of ice-cold 1X D-PBS. Cells were lysed on ice directly in plate by the addition of 2 mL of IP Lysis buffer to one 20 mm x 150 mm Cell^{+} tissue culture dishes. Cells were scraped into the lysis buffer using a chilled rubber-bladed cell scraper (Sarstedt). Cells were sheared by pipetting up and down using a pipette tip and then transferred to the next washed 20 mm x 150 mm dish of washed cells. The cells were then scraped into the same 2 mL volume of ice-cold lysis buffer transferred from the first plates. Cells were scraped into the lysis buffer and sheared by pipetting in the same manner as for the first plate. This procedure was repeated for the first 5 dishes to be harvest. For the second set of 5 dishes an additional 2 mL of ice-cold lysis buffer was added and cells were scraped and sheared in the same way as the first five dishes. The entire 4-5 mL of cell lysate was transferred to ice-cold 2 mL round-bottom tubes (Sarstedt) and lysis was allowed to proceed on ice for 30 minutes. Lysates were clarified by centrifugation of lysates at 14, 000 x g for 20 minutes at 4°C. Clarified whole cell lysate was transferred to fresh 2 mL tubes containing 50 µL of Sepharose CL-4B (Sigma) beads pre-equilibrated in cell lysis buffer. Lysates were incubated with Sepharose CL-4B resin for 45 minutes at 4°C with nutation. Following this pre-clearing step, tubes were centrifuged at 1, 000 x g for 5 minutes at 4°C to pellet the Sepharose CL-4B resin. Pre-cleared supernatants were transferred to another ice-cold

2.0 mL round-bottom tube containing 30 µL of anti-FLAG® M2 agarose (Sigma) or anti-HA High Affinity Matrix (Roche). Lysates were incubated with anti-FLAG® M2 agarose or anti-HA High Affinity Matrix for 3 hours with nutation. Samples were centrifuged at 1000 x g for 5 minutes at 4°C and supernatants discarded. The anti-FLAG® M2 agarose and anti-HA High Affinity Matrix beads were washed by resuspending the beads in 1 mL of IP Wash buffer 1, transferring the resuspended beads into a new, clean 1.5 mL tube (Eppendorf) followed by centrifugation 1000 x g for 5 minutes at 4°C. In this manner, resins were washed 3 times in IP Wash buffer 1 and once in IP Wash buffer 2. At each wash step, the resins were transferred to a new, clean, pre-chilled tube. After the 4th and final wash, resins were snap frozen on dry ice and stored at -80°C until SDS-PAGE was to be performed.

2.2.3.8 Immunopurification of endogenous proteins from mESCs for LC-MS/MS

The protocol used is the same as described in Section 2.2.3.7 with the exception that five 20 mm x 150 mm Cell^{+} tissue culture dishes (Sarstedt) were harvested per IP and the affinity resins used. For endogenous IPs, antibodies recognizing epitopes presented by the native protein were conjugated to Protein A/G Plus agarose (Pierce). Affinity resins targeting the endogenous Esrrβ and Sall4 proteins, the monoclonal anti-Esrrβ antibody (Perseus Proteomics) and the polyclonal anti-Sall4 antibody (LifeSpan Biotechnologies) were used. These antibodies were coupled to Protein A/G Plus agarose using the Protein A/G agarose IP kit (Pierce) and following the procedures exactly as outlined in the manufacturer's instruction.

2.2.3.9 Immunopurification of epitope-tagged proteins from mESCs for NativePAGE and NativePAGE-SDS-PAGE

The protocol used here is the same as the protocol described in Section 2.2.3.8 with the exception that protein complexes were eluted using 0.2 µg/mL 3XFLAG® peptide (Sigma) in IP Wash buffer 2. In brief, 3XFLAG-Nanog was eluted by resuspending the anti-FLAG® M2 agarose in 500 µL of 0.2 µg/mL 3XFLAG® peptide solution. These elution mixtures were placed in an Eppendorf Thermo mixer set to 2500 rpm for 25 minutes at RT. This elution was repeated one more time. Eluates were pooled together and concentrated by centrifugation using an Amicon® Ultra-0.5mL, 3K MWCO device (Millipore).

2.2.3.10 λ-phosphatase treatment of immunopurified proteins for SDS-PAGE

3XFLAG-Nanog stable cell line was grown in 10 cm Cell^{+} dishes (Sarstedt) and lysed in 1mL of IP lysis buffer containing as per protocol described in Section 2.2.3.7 Immunopurification of epitope-tagged proteins from mESCs for LC-MS/MS. Following final wash, pelleted resins were treated with either λ-phosphatase (NEB) buffer containing λ-phosphatase and $MgCl_2$, λ-phosphatase and phosphatase inhibitors or no λ-phosphatase (UT Control) as per manufacturer's instructions. Tubes were vortexed for 1 minute every 10 minutes throughout the incubation. The reactions were stopped by adding 50 µL of 4X NuPAGE® LDS sample loading buffer containing 100µM DTT and boiling for 10 minutes at 100°C.

2.2.4 Mass spectrometry methods:

2.2.4.1 SDS-PAGE for mass spectrometry

The protocol is the same as described in Section 2.2.3.2 with the following exceptions: The XCell SureLock™ Electrophoresis Cell used for SDS-PAGE for LC-MS/MS was cleaned with ethanol, rinsed with ddH_2O filtered through a 0.2 µm membrane and allowed to air-dry thoroughly before use. All sample loading buffers used resuspend resins for sample loading were filtered through 0.2 µm membrane. After electrophoresis, gels were placed in a plastic 15 cm x 15 cm plastic petri dish filled with 0.2 µm membrane-filtered ddH_2O and washed for 15 minutes at RT with orbital shaking. This wash was repeated two more times before decanting the water and adding 25 mL of SimplyBlue™ SafeStain (Invitrogen) Coomassie stain compatible with mass spectrometry. Gels were incubated in the Coomassie stain for 1 hour at RT with orbital shaking. The stain was removed and 30 mL of 0.2 µm membrane-filtered ddH_2O containing 2% (w/v) NaCl was added. Gels were incubated in this destaining solution overnight at RT with orbital shaking. The following morning, the destaining solution was removed and replaced with fresh 0.2 µm membrane-filtered ddH_2O and allowed to destain for another hour at RT with orbital shaking prior to scanning the gel.

2.2.4.2 Gel band cutting

Working in a laminar flow hood, gels were placed on a clean Niceday transparency (Office Depot) and the entire sample lane was excised as 5 or 6 large blocks using a sterile scalpel. Gel slices were transferred into sterile 2.0 mL round-

bottom tubes (Eppendorf), large gel slices were further chopped into smaller pieces stored at 4°C until digestion with trypsin was to be performed.

2.2.4.3 In-Gel digest

Gel pieces were shrunk by adding 200 µL of methanol to the tubes and incubating for 10 minutes at RT. Methanol was removed and 200 µL of 50 mM ammonium bicarbonate solution was added for 10 minutes at RT. Alternating incubation in methanol and 50 mM ammonium bicarbonate solution was repeated two more times. Gel slices were then reduced by adding 100 µL of 20 mM DTT in 50 mM ammonium bicarbonate solution. Gel slices were incubated in this solution for 30 minutes at 60°C. The DTT-containing solution was removed and 200 µL of methanol was added for 10 minutes at RT. Proteins in the gel slices were then alkylated by adding 100 µL of 50 mM iodoacetamide in 50 mM ammonium bicarbonate solution. Gel slices were incubated in this solution for 1 hour at RT. The iodoacetamide-containing solution was removed and 200 µL of methanol was added. Gel slices were incubated in methanol for 10 minutes at RT or until the gel pieces had shrunken. All traces of methanol were removed and gel slices were allowed to dry at RT. To the dried gel slices, 20 µL of trypsin solution [20 nM] was added and gel slices were incubated in this solution for 30 minutes at 4°C. Enough [50 mM] ammonium bicarbonate solution was added to cover the gel slices and gel slices were allowed to digest overnight at RT. The following day, gel digest solutions were transferred to a new clean tube and 50 µL of 1% (w/v) formic acid solution was added to the sample. Samples were incubated for 10 minutes at RT. The solution extracts were then transferred to the previous tube and 50 µL of 1% formic acid in (50:50)

water/methanol solution was added to the samples. After 15 minutes of incubation at RT, the extracts were transferred back to the previous tube and the previous two steps were repeated another time. Finally, the solution extracts were dried and concentrated using a Speedvac (Thermo Fisher).

2.2.4.4 LC-MS/MS analysis protocol

Samples were analyzed using an on-line system consisting of a micro-pump (1200 binary HPLC system, Agilent, UK) coupled to a hybrid LTQ-Orbitrap XL mass spectrometer (Thermo Fisher, UK). The LTQ was automatically controlled using Xcaliber 2.0.7 software. Samples were injected into the LTQ-Orbitrap XL by capillary picotip columns (10 cm x 360 µm OD x 75 µm ID) with a 15 um tip opening size and fitted with borosilicate frits (New Objective, PResearch, UK). Pre-columns (4 cm x 360 µm OD X 200 µm ID) were used in a vented column pre-column arrangement (Licklider *et al.*, 2002; Le Bihan *et al.*, 2003). Fused silica tubing was purchased from Composite Metal (UK). Digested samples were resuspended in 10 µL of reverse phase loading buffer containing 95.7% H_2O, 2.5% acetonitrile and 0.1% (w/v) formic acid. Eight µL was injected into the HPLC at a flow rate of 5 µL/min. After sample loading, the flow rate across the injection column was reduced to approximately 100-200 nL/min using the vented column arrangement. The peptide mixture was separated by a C18 reverse-phase column using a 2 hour elution gradient from 0% - 100% of buffer composed of 90% acetonitrile, 10% (v/v) H_2O, 0.025% trifluoroacetic acid (w/v) and 0.1% (w/v) formic acid. The mass spectrometer was operated in data-dependant mode, with a single MS

scan (400-2000 m/z) in FT mode 60K resolution followed by MS/MS scans in the linear ion trap on the 5 most abundant ions and dynamic excluded for 120 seconds.

2.2.4.5 Data analysis with MaxQuant software

Data acquired from the MS/MS runs were searched against a mouse plus contaminant IPI database containing 55413 sequences downloaded from www.ebi.ac.uk using MASCOT v 2.3 software (Matrix Science Ltd., UK).

All MS/MS data acquired in this study were searched for the following post-translational modifications: variable methionine oxidation, STY phosphorylation, protein N-terminal acetylation and fixed cysteine carbamidomethylation. The precursor mass tolerance was set to 7ppm and MSMS tolerance to 0.4amu. The significance threshold (p) was set below 0.05 (MudPIT scoring in Mascot). All LC-MS runs performed in this study were combined using MaxQuant v 1.0.13.8 assuming a false positive rate of 0.01 as per protocol described by Cox *et al.*, 2009. The MaxQuant output of protein group intensities was calculated as the sum of MS-peak intensities of identified peptides. This sum was used as a measure of relative abundance between samples.

In these studies, a high confidence interactor was defined as a protein for which at least 2 unique peptides were identified. Where the number of peptides detected in the sample was 4-fold greater than in the control as well as a peptide peak intensity ratio 20-fold greater than in the control as per Lowery *et al.*, 2007.

2.2.4.6 Protein network diagram generation using STRING v 9.0 and Cytoscape version 2.8.1

The list of high confidence interacting proteins generated by the MaxQuant software output, were analyzed for previously reported protein-protein interactions using the platform tool STRING v 9.0 (http://www.string-db.org). Known interactors were identified using the *mus musculus* database and selecting the experimental evidence, database and yeast two-hybrid search parameters and, importantly, de-selecting the default co-occurence criteria. The STRING v 9.0 med-high confidence scoring criteria was used to filter the interactions reported.

2.2.5 Immunofluorescence methods:

2.2.5.1 Immunostaining of mESCs for quantitative confocal microscopy

Cells were grown on gelatin-coated 8-well Lab-Tek® II Chamber Slide™ (ThermoFisher). Growth medium was removed from each well of the 8-well chamber slide to be immunostained. Chamber slides washed with 500 µL of room temperature 1X PBS (Invitrogen) once for 5 minutes. Cells were fixed by adding 500 µL of IF fixing solution and incubating for 15 minutes at RT. Cells were washed 3 times for 5 minutes with 500 µL of 1X PBS. Cells were permeabilized by 500 µL of IF permeabilization solution and incubating for 15 minutes at RT. Blocking was performed by the addition of 500 µL of IF blocking solution for 1 hour at RT. Primary antibodies were diluted (1:100) in IF blocking solution. After 1 hour blocking at RT, the blocking solution was removed and 250 µL of primary antibody solution and incubated overnight at 4°C. Cells were then washed 3 times for 5 minutes with IF blocking solution and 500 µL of secondary antibody solution was added to each well. All secondary antibody solutions were prepared as [1:500]

dilutions in IF blocking solution. Cells were incubated with secondary antibody solutions for 1 hour at RT in the dark. Antibody solutions were then removed and cells were washed 3 times for 5 minutes in 1X PBS solution at RT. DNA was stained by adding 500 µL of IF DAPI staining solution to each well and incubating at RT for 10 minutes in the dark. The IF DAPI staining solution was removed and slides were mounted by the addition of one drop of ProLong® Gold antifade reagent (Invitrogen) and gently overlaying a 24mm x 60mm rectangular coverglass (Corning) over all 8 wells. Slides were stored protected from light until ready to be analyzed by confocal microscopy.

2.2.5.2 ROS Assay

Cells were grown on gelatin-coated 8-well Lab-Tek® II Chamber Slide™ (ThermoFisher). Growth medium was removed and replaced with either regular growth medium or fresh medium containing the CellROX™ Deep Red Reagent (Invitrogen) to a final concentration of 1 µM. Chamber slides with fresh medium containing dye were returned to standard culture conditions or hypoxic culture conditions for 30 minutes at 37°C. Medium was then removed from each well of the chamber slide and washed with 500 µL of room temperature 1X PBS (Invitrogen) once for 5 minutes. Cells were fixed by adding 500 µL of IF fixing solution and incubating for 15 minutes at RT protected from light. Cells were washed 3 times for 5 minutes with 500 µL of 1X PBS. In these studies, the ROS assay was performed as one part of a multiplexed fluorescence-based assay for quantitation by confocal microscopy. Therefore, after performing the three final washes of post cell permeabilization, 500 µL of IF blocking solution was added to all wells. Subsequent

immunostaining and mounting procedures were performed as described in Section 2.2.5.1.

2.2.5.3 Immunostaining of mESCs in hypoxic conditions

The protocols used were the same as those described in Sections 2.2.5.1 - 2.2.5.2 with the following exceptions: Prior to cell fixation and permeabilization, all work was carried out in the XVivo model G300C hypoxia workstation under hypoxic conditions defined above. All growth medium, 1X PBS and cell fixation solutions were placed in the workstation with lids off for a minimum of 4 hours prior to beginning immunostaining procedures. All drug treatments and the CellROX™ Deep Red Reagent was added to cells in hypoxic conditions. Following cell fixation and permeabilization under hypoxia, all chamber slides were removed from the XVivo workstation and the remainder of the procedure was performed in normoxic conditions.

2.2.5.4 Quantitative confocal microscopy image acquisition and data extraction

All confocal microscopic images were acquired using the Leica SP5, Model TCS SP5 confocal microscope mounted on the SP5 workstation (Leica Microsystems CMS, GmbH) located in the C.O.I.L. The microscope and image acquisition parameters were controlled automatically using LAS AF Software program (Leica). All microscopy in this study was performed by acquiring 3 z-stacks having a z-step size of 0.29 µm and through a minimum z-volume of 12 – 17 µm (depending on colony size and height). In xyz acquisition mode, images were acquired in 1024 x 1024 format at a speed of 400 Hz, and 12-bit bit depth. The size of the images

obtained was 246.0 µm x 246.03 µm having a pixel size of 240.50 nm x 240.50 nm x 293.73 nm. Images were obtained at a 630X total magnification using HCX PL APO CS 63.0 x 1.40 objective and the Argon (DAPI), DPSS 561nm (Alexa®488), HeNe 594nm (Alexa®555 and Alexa®594) and HeNe 633nm (Alexa®647) lasers.

Data was extracted from images by three methods:

1) Using ImagePro v 7.0 software along with Macro1 written by Dave Kelly – C.O.I.L. facility manager (See APPENDIX F). Briefly, this Macro measured the green and red intensities based on the nuclear outlines produced by the thresholding of the DAPI image. The DAPI outlines were used to remove the nucleus from the red image. Once the nuclear signal was removed, the rest of the red cell had its intensity measured.

2) Using ImagePro v 7.0 along with Macro3 written by Dave Kelly – C.O.I.L. facility manager (see APPENDIX G). Briefly, this Macro quantified the intensity in the green channel and then counted fluorescent foci in the red and far red channels. The DAPI channel was thresholded by the user and then used to map the measurements made in the other channels to the nucleus. The area of the nucleus was also measured.

3) Using Volocity® v 6.0 software (Perkin Elmer, UK). Briefly, this software created outlines of the DAPI signal and used these outlines to map the quantified intensities of all 4 channels. In addition, this software measured nuclear area and counted fluorescent foci both in the nucleus and whole cell area.

2.2.5.5 Co-localization analysis using Volocity® 6.0 software.

Using Volocity® (Perkin Elmer) v 6.0 software package, unthresholded co-localization statistics, Pearson's co-efficients, as well as, overlapping pixels images were generated for all channels of interest from one representative z-stack obtained at 0.29 µm step size through the entire cell depth (12-20 µm).

CHAPTER 3: Pluripotency Factor Protein-Protein Interaction Mapping

3.1 Introduction

Mouse embryonic stem cells (mESCs) possess both the ability for unlimited self-renewal in culture and the capacity to maintain their pluripotency indefinitely (Smith *et al.*, 2001). These unique abilities are regulated, in part, by the transcription factors Nanog, Oct4 and Sox2 (Niwa *et al.*, 2007).

Recent studies have shown that Nanog is post-translationally regulated both by phosphorylation which stabilizes it and ubiquitination which targets the protein for degradation by the 26S proteasome (Moretto-Zita *et al.*, 2010; Ramakrishna *et al.*, 2011; Chae *et al.*, 2012). Therefore, the development of a strategy to identify those proteins involved in the post-translational modification of Nanog and other pluripotency factors would provide valuable insight into how pluripotency in regulated in ES cells.

To date, several IP-LC-MS/MS protein-protein interaction studies of Nanog and other key pluripotency factors in mESCs have been published (Wang *et al.*, 2006; Van den Berg *et al.*, 2008; Liang *et al.*, 2008; Van den Berg *et al.*, 2010; Mak *et al.*, 2010; Pardo *et al.*, 2010; Mallanna *et al.*, 2010). Each of these protein interaction network datasets were obtained using different cell types, starting material (nuclear versus whole cell extracts) and purification methodologies. The integration of these networks using Cytoscape reveals that the 10 different pluripotency factors are highly interconnected Mallano and Rizzino, 2012.

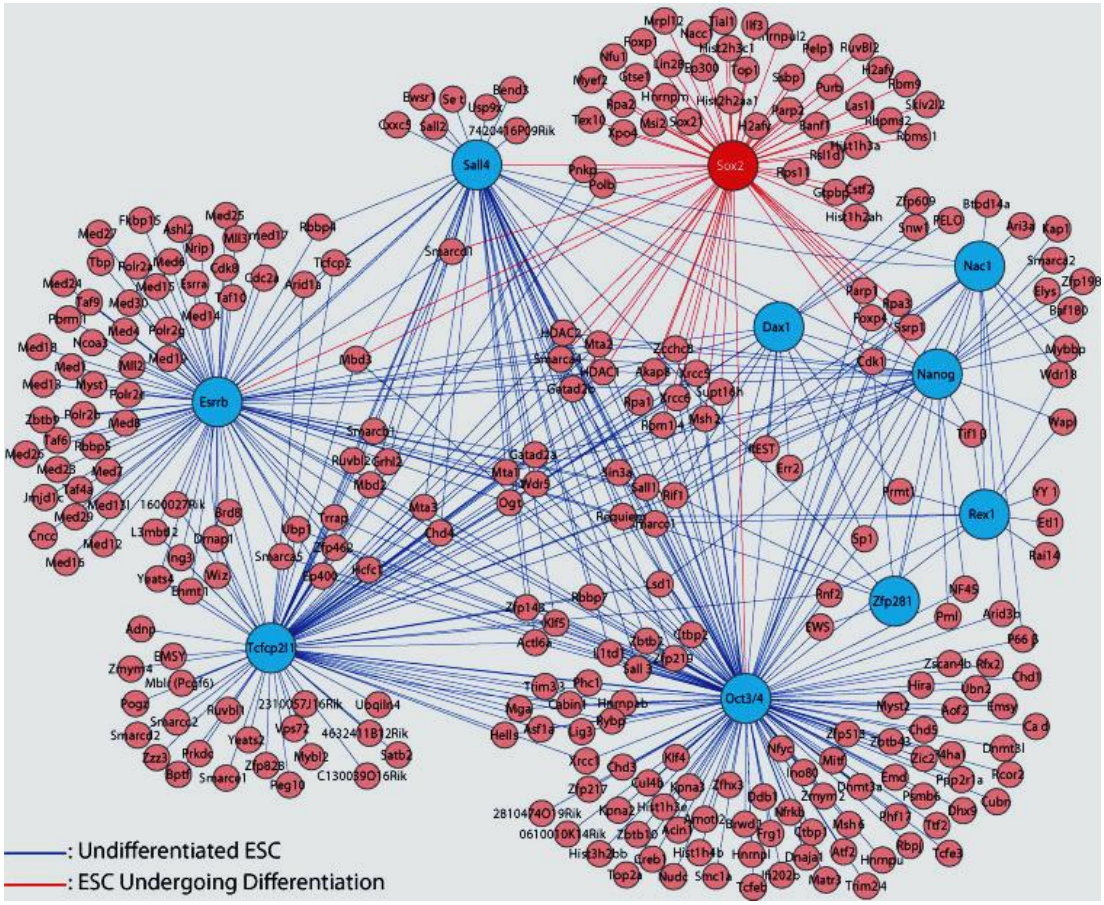


Image taken from Mallanna and Rizzino, *J Cell Physiology* 2012.

Figure 3.1 The published protein–protein interaction datasets for 10 different pluripotency factors obtained from mESCs. Using Cytoscape, the IP-LC-MS/MS datasets for 10 different transcription factors were merged to generate an integrated protein interaction network for mESCs. The interactors identified for Nanog, Oct4, Esrrβ, Sall4, Rex1, Dax1, Tcfcp1l1, Zfp281 and Nacc1 in mESCs (coloured in blue) (Wang *et al.*,2006; Liang *et al.*, 2008; Pardo *et al.*, 2010; Van den Berg *et al.*, 2010) were integrated with the interactors identified for Sox2 undergoing early-stage differentiation (coloured in red) (Mallanna *et al.*, 2010).

Although there are few direct interactions reported between the pluripotency factors themselves (Figure 3.1), they do interact with many of the same proteins. Furthermore, several of these shared interactors between pluripotency factors combine to form complexes which further connect the various interaction networks to each other. This suggests that these 10 essential pluripotency factor proteins are highly interconnected and, therefore, must be tightly regulated.

Proteins and protein function are regulated in several ways including, but not limited to, phosphorylation and ubiquitination. In recent years, Nanog, Oct4 and Sox2 have been shown to have their functions modulated by phosphorylation (Moretto-Zita *et al.*, 2010; Seki *et al.*, 2010; Jeong *et al.*, 2010). A protein-protein interaction network obtained from whole cell extracts would be necessary in order to identify the components of phosphorylation-mediated signaling cascades feeding into the modulation of these master regulators.

Another reason why whole cell lysates should be used in IP-LC-MS/MS studies is due to the characteristic physical properties of ESCs. That is, their high nucleus to cytoplasm ratio (Smith 1992). ESCs having a large nucleus surrounded by a small amount of cytoplasm would, in theory, allow for signaling cascades to be transduced more rapidly in these cells compared to their differentiated counterparts. It follows that the protein interaction landscape described by Mallanna and Rizzino (2012) (Figure 3.1) would be tightly regulated by finely tuned balances between kinases and phosphatases, ubiquitin ligases and peptidases, acetyltransferases and deacetylases, etc. Since many of these reactions may occur within the cytoplasm, it should be included in order to identify those proteins involved in the post-translational modification of pluripotency factors.

Only one of the protein-protein interaction networks published to date (Figure 3.1) has been performed on whole cell lysates (Pardo *et al.*, 2010). The remaining studies have all been performed on nuclear extracts. This is presumably because these pluripotency factors are transcription factors and transcription factors are usually found in the nucleus. However, there is a long list of transcription factors that are known to shuttle between the nucleus and the cytoplasm. For example, Bach1

(Yamasaki *et al.*, 2005), FoxO1 (Gan *et al.*, 2005), NF-κB (Mikenberg *et al.*, 2007) and the pluripotency factors Oct4B (Liedtke *et al.*, 2007) and Sox2 (Baltus *et al.*, 2009; Zhao *et al.*, 2011)

In the study performed by Baltus *et al.*, 2009 it was demonstrated that the acetylation of Sox2 results in its translocation to the cytoplasm where it is ubiquitinated and targeted for degradation. In light of these findings, elucidating the Sox2 protein-protein interaction network using whole cell lysates as opposed to nuclear extracts could provide valuable insight into the post-translational regulation of this protein.

Due to the fact that several transcription factors are known to have a life outside the nucleus characterizing a protein interaction network under limiting conditions such as nuclear extracts although highly specified would provide a deep but potentially narrow view of the pluriproteomic landscape.

In this study, two pluripotency factors known to be modulated by post-translational regulation were chosen as baits: Nanog and Sox2. Sall4 was chosen due to its reported interaction with Nanog (Wu *et al.*, 2006; Zhang *et al.*, 2006). Esrrβ and Pik3cd were identified as interactors with Nanog in this study and were chosen in order to complete the network.

3.2 Aims of this chapter

The aim of this chapter is to elucidate a protein-protein interaction network centered around Nanog and Sox2 that extends beyond the nucleus of mouse embryonic stem cells. This will be accomplished by immunopurifying Nanog and Sox2 from whole cell lysates of mESCs and subsequently performing LC-MS/MS

analysis on the isolated protein complexes. Additionally, two interactors identified in the Nanog purifications will be chosen in order to reciprocate parts of the network. The data will be analyzed using STRING and Cytoscape in order to identify the connections within the network and uncover those biochemical pathways linking extracellular and cytoplasmic signaling events that feed into the regulation of pluripotency in mESCs. One further goal of this chapter is to identify those proteins involved in post-translational modifications that interact with these pluripotency factors. The identification of either the kinase(s) responsible for phosphorylating Nanog or the protein(s) that target it for degradation by the Ubiquitin-Proteasome System would provide valuable insight into the molecular mechanisms by which pluripotency is maintained.

3.3 Results:

3.3.1 Nanog protein-protein interaction network

In order to obtain a protein-protein interaction map for Nanog, a mESC cell line stably expressing 3XFLAG epitope-tagged mouse Nanog protein was first established. As shown in Figure 3.2, 3XFLAG-Nanog protein complexes were immunopurified from whole cell lysates prepared from the E14tg2a/3XFLAG-Nanog mESC cell line using anti-FLAG M2 agarose. In order to eliminate any potential artifacts introduced due to the use of the puromycin selection antibiotic, an anti-FLAG M2 agarose IP performed on whole cell lysates prepared from TNG cells similarly cultured in 1µg/mL puromycin was used as the control. The entire sample lanes were cut and digested ensuring that bias due to differential Coomassie staining was not introduced into the results. The results are based on 2 biological replicates of anti-FLAG M2 agarose IPs from 3XFLAG-Nanog whole cell lysates compared to 2

biological replicates of anti-FLAG M2 agarose IPs from TNG whole cell lysates (Figure 3.2).

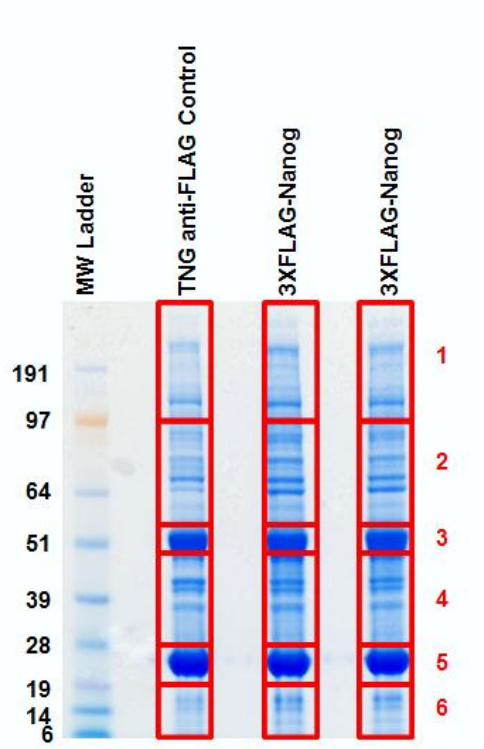
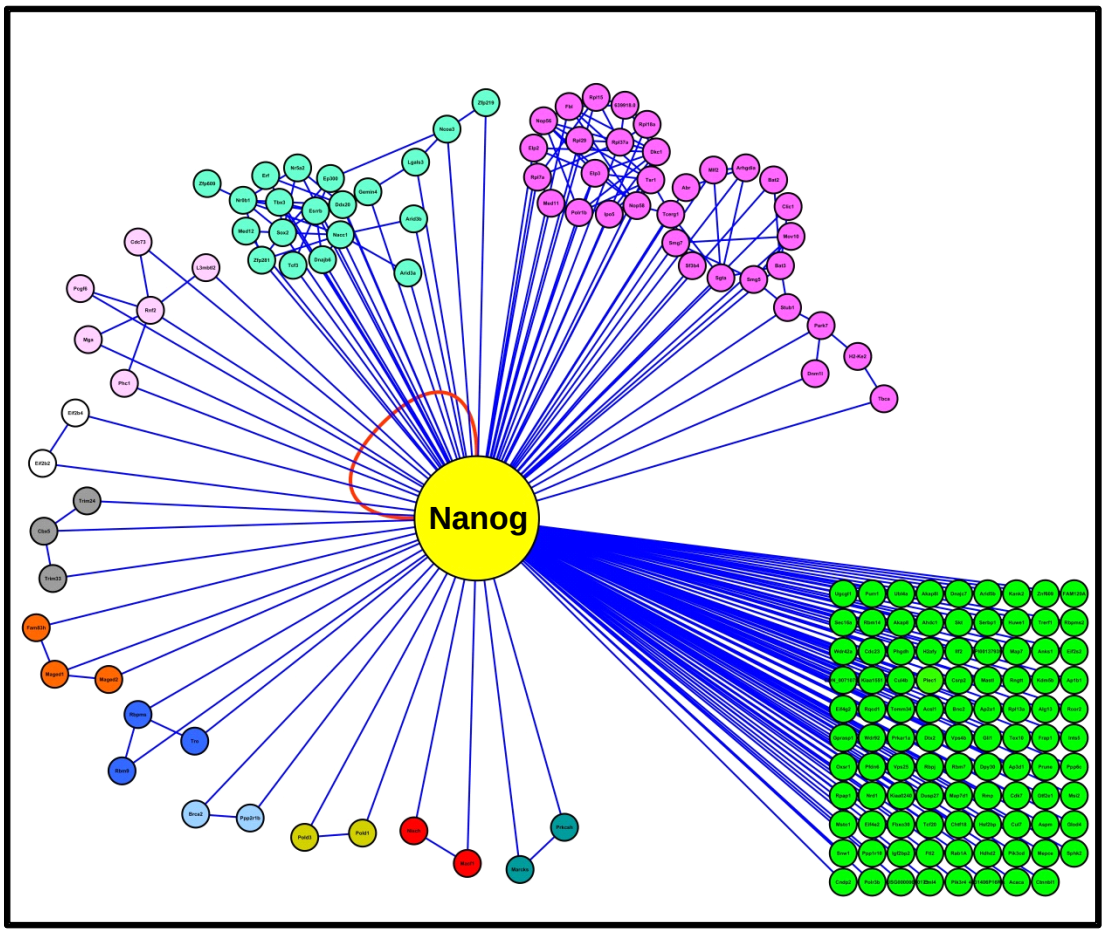


Figure 3.2: Immunopurification of 3XFLAG-Nanog protein complexes. Coomassie stained gel of 3XFLAG-Nanog protein complexes isolated using anti- FLAG M2 agarose. Gel slices excised for digestion and subsequent LC-MS/MS analysis are indicated by red boxes and numbers.

In order to obtain a protein-protein interaction network consisting of nodes (proteins) and edges (interactions) for the Nanog protein in mESCs, the LC-MS/MS data acquired from these experiments was analyzed by searching for known interactions between the proteins identified using STRING and were graphically illustrated using Cytoscape (Figure 3.3).



RNA transcription, elongation and splicing
Nuclear hormone receptors and Sox2-containing cluster
Epigenetic transcriptional repression complex
Translation initiation/elongation
Heterochromatin associated
Apoptosis signaling
RNA splicing
BRCA2-containing complex
Transcription initiation/elongation
Cytoskeleton remodeling
PKC signaling
Remaining Nanog interactors (not part of the 11 clusters)

Figure 3.3: The Nanog protein-protein interaction network generated in this study. Cytoscape diagram of the high-confidence Nanog interacting proteins identified by LC-MS/MS. Proteins that are reported to interact with each other are clustered together and differentially coloured based on the known function of these interacting groups as reported by STRING software. Each of the protein clusters is further illustrated in Figures 3.4 and 3.5.

In these experiments, 174 high confidence interactors were identified for 3XFLAG-Nanog and are listed in APPENDIX A. In figure 3.3, the Nanog interaction network is shown having 174 nodes and 273 edges. Of the 174 interacting proteins identified, 75 were clustered into 11 interacting protein groups having known functions. The interacting protein clusters were coloured based on their known function ascertained from the STRING protein descriptions output and the bright green-coloured nodes represent the other 99 proteins identified to interact with Nanog. The circular red line on the Nanog node indicates that this protein interacts with itself. This interaction has been well characterized as Nanog self-dimerization (Mullin *et al.*, 2008) and is essential for the maintenance of pluripotency (Wang *et al.*, 2008).

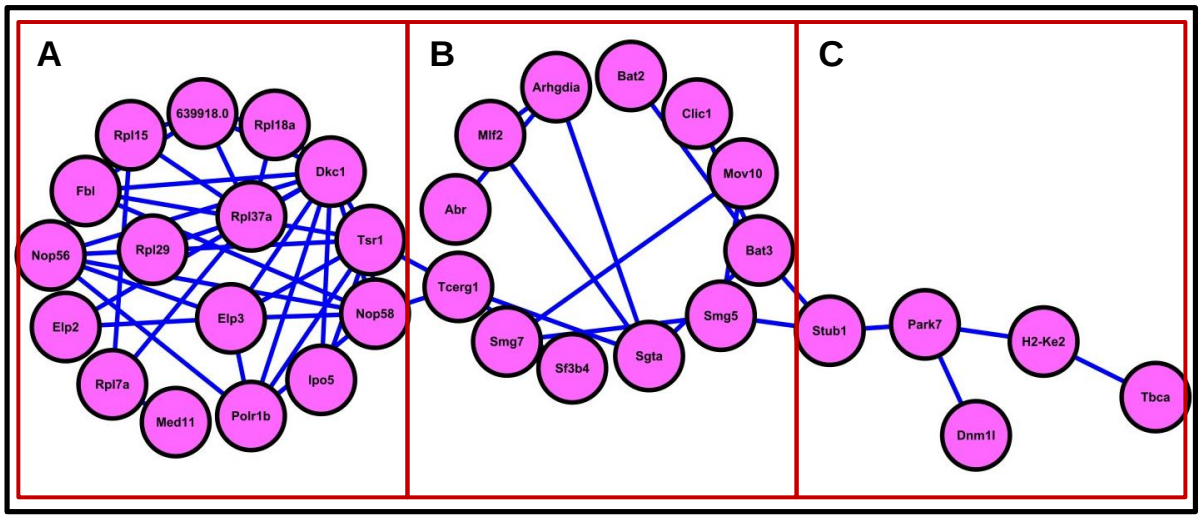


Figure 3.4: The mRNA transcription initiation, elongation and splicing cluster. Cytoscape diagram of the largest cluster of proteins found to interact with Nanog. **A)** Subgroup of proteins linking mRNA transcription initiation with transcript synthesis, **B)** Subgroup of proteins involved in mRNA splicing and metabolism, **C)** Subgroup of oxidative stress sensing and response proteins.

The largest interacting cluster of proteins identified is shown in Figure 3.4. This large cluster consisting of 33 proteins can be further sub-divided into 3 subgroups. The first subgroup, shown in Figure 3.4 (A), consists of proteins involved in mRNA transcript initiation and transcript synthesis/elongation. The proteins Tsr1

and Nop58 provide potential links between this subgroup and the mRNA splicing subgroup shown in Figure 3.4 (B) Smg5 from the mRNA splicing and metabolism subgroup links, via its interaction with Stub1(Lee *et al.*, 2006), to a subgroup of proteins involved in oxidative stress sensing and response.

The remaining 10 interaction clusters identified in this study are shown in Figure 3.5 A(i-x). A(i) Shows a large cluster of proteins containing the ligand-dependent nuclear hormone receptors Esrrβ, Nr5a2 and Nr0B1. Also in this cluster is the master pluripotency regulator Sox2 and some of its previously reported interactors, Zfp281 and Nacc1. A(ii) Shows a group of proteins involved in epigenetic-mediated gene silencing. They are, Rnf2, Phc1 and Cdc73. A(iii) Shows the translation initiation proteins Eif2b2 and Eif2b4. A(iv) Shows a small cluster of proteins containing Cbx3, Trim24 and Trim33 which are known to be associated with heterochromatin. A(v) Shows the Fam83h, Maged1 and Maged2-containing cluster of proteins involved in apoptotic signaling. A(vi) Shows a cluster of additional mRNA splicing proteins. A(vii) Shows the novel interactors BRCA2 and Ppp2r1b. A(viii) Shows the Pold1 and Pold3 proteins involved in the synthesis of DNA. A(ix) Shows a pair of proteins (Nisch and Macf1) thought to be involved in cytoskeleton remodeling. A(x) Shows a cluster of proteins containing substrates for protein kinase C signaling. Namely, they are: Prkcsh (protein kinase C substrate 80K-H) and Marcks (myristoylated alanine-rich protein kinase C substrate).

In order to confirm some of the Nanog interacting proteins observed by mass spectrometry, a series of co-immunopreciptation experiments were performed. In Figure 3.5(B), a western blot performed on protein complexes isolated using anti-FLAG M2 agarose beads from the 3XFLAG-Nanog mES cell line and TNG cell line

is shown. The membrane was probed with antibodies recognizing the endogenous Esrrβ, Nr0B1 and Nanog proteins.

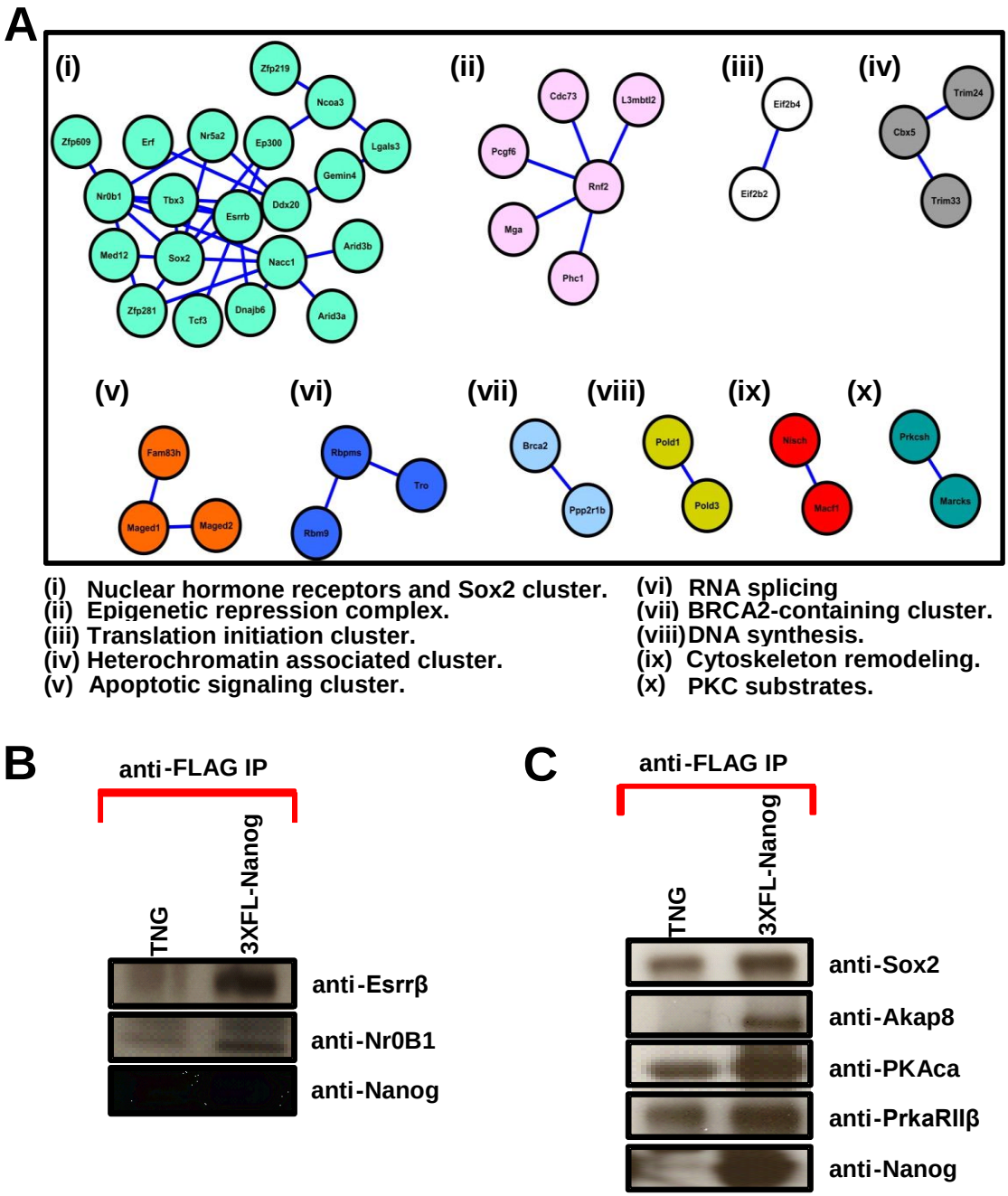


Figure 3.5: **The remaining 10 Nanog interaction clusters and co-immunoprecipitation of 3XFLAG-Nanog with several identified interactors. (A)** Cytoscape diagram of the Nanog interacting protein clusters differentially coloured based on the reported functions of these interactors using STRING. **(B)** Western blotting performed on anti-FLAG immunopreciptation of 3XFLAG-Nanog complexes from mESCs probed with antibodies recognizing the endogenous co-purifying nuclear hormone receptors Esrrβ and Nr0B1/Dax1. **(C)** Western blotting performed on anti-FLAG immunopreciptation of 3XFLAG-Nanog complexes from mESCs probed with antibodies recognizing the endogenous co-purifying Sox2 as well as several components of the PKA signaling pathway.

In Figure 3.5(C), a western blot performed on protein complexes isolated using anti-FLAG M2 agarose beads from the 3XFLAG-Nanog mES cell line and TNG cell line is shown. The membrane was probed with antibodies recognizing the endogenous Sox2, Akap8, PKACa, PrkaRIIβ and Nanog proteins.

Several proteins involved in performing various forms of post-translational modifications were identified as part of the Nanog protein interaction network. In total, 6 kinases, 2 cullins, 5 E3 ubiquitin ligases and 1 F-box protein were identified by IP-LC-MS/MS. These protein names, number of peptides identified and peptide intensities are summarized in Table 3.1.

Protein Class	Gene Name	Number of Peptides Detected	% Sequence Coverage	Intensity Ratio (sample/control)
Kinase	Pik3cd	2	2	*
	FRAP1 (mTOR)	3	1	*
	Sphk2	2	3	*
	Mastl	3	6	*
	Oxsr1	2	4	*
	Prkar1a	2	6	*
	Cdk7	2	5	*
Cullins	CUL4B	7	**7**	39
	CUL7	3	1	*
E3 Ub Ligase	Trim33	6	6	*
	Rnf2	8	32	31
	Huwe1	28	8	*
	Dtx2	3	8	*
	Cdc23	5	8	21
F-box Protein	Fbxo30	2	3	*

*** Indicates a value of 0 was detected in the control.**
Yellow box indicates that all peptides detected are unique to the CUL4B isoform.

Table 3.1 Nanog interacting proteins involved in post-translational modifications based on n=2 experiments. Summarized are the proteins identified by IP-LC-MS/MS in this study to co-purify with Nanog protein. The intensity ratio (sample/control) reported here is an intensity measure using protein intensity extracted by MaxQuant software which corresponds to the sum of the intensity observed for the MS/MS-identified peptides composing the protein.

3.3.2 Sox2 protein-protein interaction network

In this set of experiments, the bait protein was the transcription factor Sox2 having six repeats of the HA epitope tag linked to its amino terminus (Maruyama *et al.*, 2005). After generating a mouse embryonic stem cell line stably expressing 6XHA-Sox2 in an E14tg2a wild-type background, 6XHA-Sox2-containing protein complexes were isolated using anti-HA High Affinity matrix, separated on an SDS-PAGE gel and stained with Coomassie (Figure 3.6). As in the previous section, the TNG cell line grown in the presence of 1µg/mL puromycin was used as control lysate. The entire sample lanes were cut from the gel and digested with trypsin for subsequent LC-MS/MS analysis.

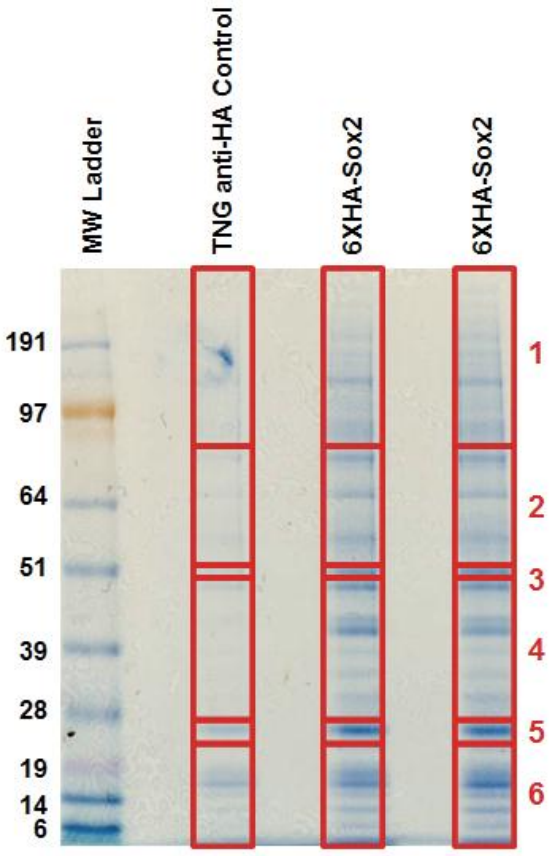


Figure 3.6: Immunopurification of 6XHA-Sox2 protein complexes. Coomassie-stained gel of 6XHA-Sox2 protein complexes isolated using anti-HA High Affinity matrix. Gel slices excised for digestion and subsequent LC-MS/MS analysis are indicated by red boxes and numbers.

The results are based on 2 biological replicates of anti-HA High Affinity Matrix IPs from 6XHA-Sox2 whole cell lysates compared to 2 biological replicates of anti-HA High Affinity Matrix IPs from TNG whole cell lysates. The protein-protein interaction network obtained for 6XHA-Sox2 in this study consisted of 145 nodes and 196 edges. The full list of Sox2-interacting proteins identified in this study is given in APPENDIX B. In Figure 3.7, 54 of the 145 proteins identified, have been clustered into 10 interacting protein groups having known functions. The interacting protein clusters were coloured based on their known function and the bright green-coloured nodes represent the other 91 proteins identified to interact with Sox2.

All 10 interaction clusters are shown in Figure 3.8 (A-J). (A) Shows a cluster of proteins containing Numa1which is a key component of the nuclear mitotic apparatus. (B) Shows a cluster containing the Nid2, Lamc1 and Hspg2 proteins involved in cell adhesion and migration. (C) Shows the Idh2 and Aprt pair of proteins involved in cellular metabolism and energy production. (D) Shows the E3 ligase Huwe1 and its putative partner Hadh which is involved in lipid metabolism. (E) Shows a large cluster of proteins containing the mitotic checkpoint proteins BUB1 and CyclinB1. Further links to the mitotic checkpoint and DNA damage repair are provided by the presence of Rad50 and ATM kinase within this group. (F) Shows a large cluster of proteins containing the proteins Polr2a and Polr2b which are involved in transcription initiation and elongation. The kinase Mapk3 (ERK1) was identified among this group of proteins, as well as, CUL3 a key component of Cullin-RING E3 ligases involved in the regulation of mitosis. (G) Shows a small cluster of proteins involved in MafK transcription activation/repression. (H) and (I) Show two clusters involved in translation initiation and elongation, respectively. (J)

Shows a pair of proteins containing Pds5a known to be involved in the segregation of mitotic chromosomes.

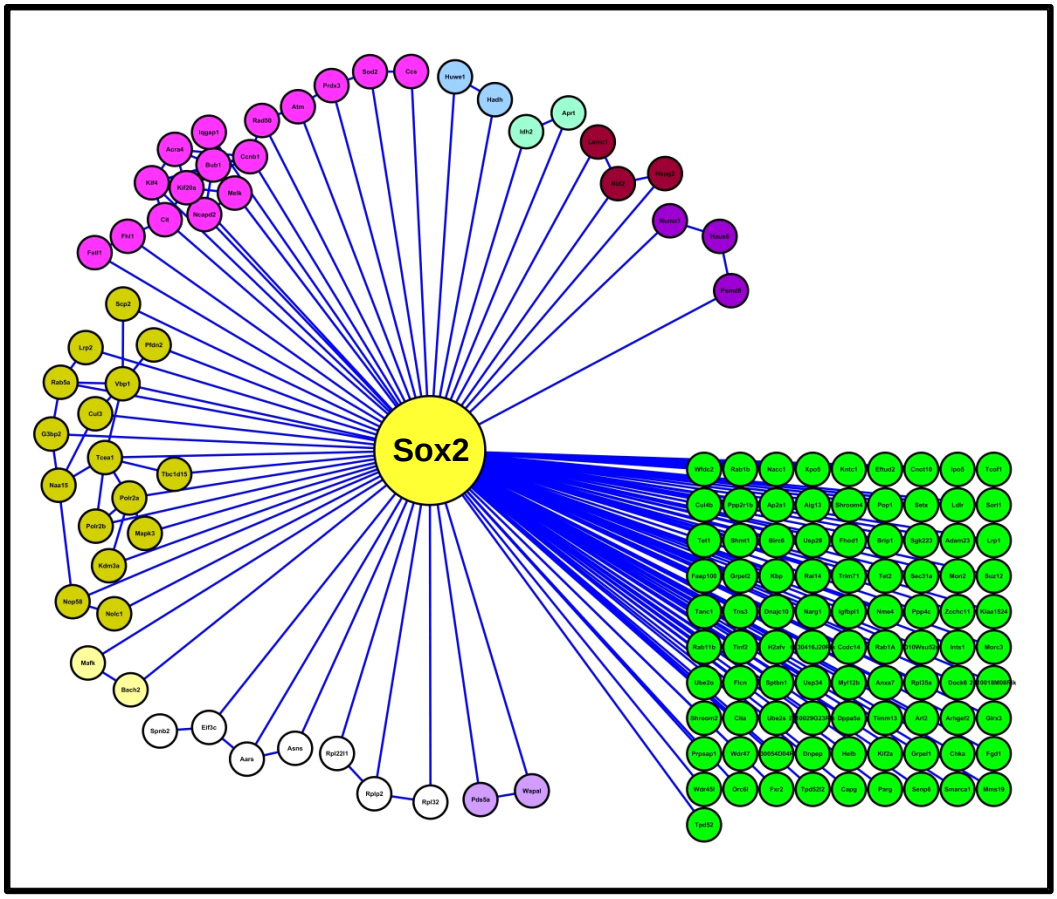
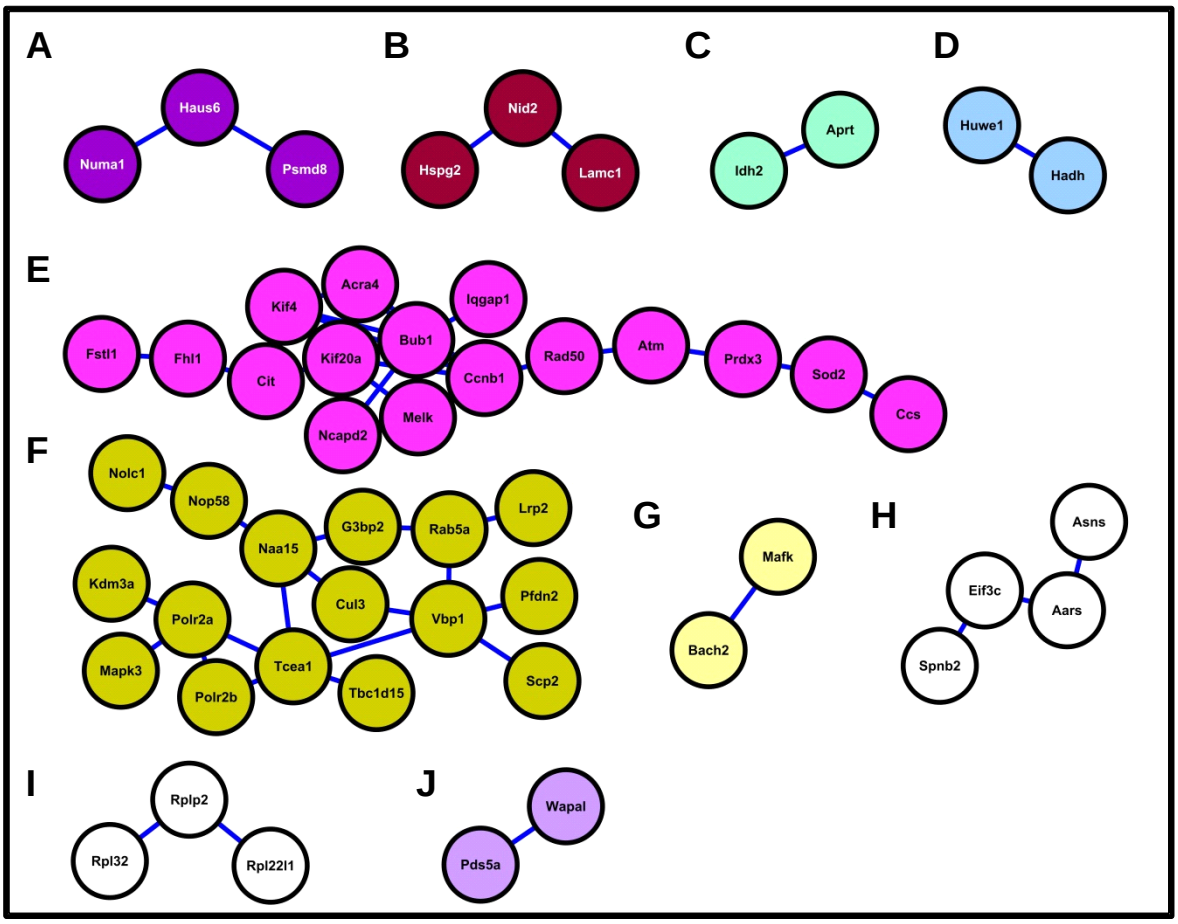


Figure 3.7: The Sox2 protein-protein interaction network generated in this study. Cytoscape diagram of the high confidence Sox2 interacting proteins identified by LC-MS/MS. Proteins that are reported to interact with each other are clustered together and differentially coloured based on the known function of these interacting groups as reported by STRING software. Each of the protein clusters are further illustrated in Figure 3.8.



A. Nuclear mitotic apparatus components.
B. Cell adhesion and migration cluster.
C. Metabolism/energy production cluster.
D. Lipid metabolism cluster.
E. Mitotic checkpoint/DNA damage repair cluster.
F. Transcription initiation/elongation cluster.
G. MafK transcription activation/repression cluster.
H. Translation initiation/elongation cluster.
I. Translation initiation/elongation cluster.
J. Chromosome segregation cluster.

Figure 3.8: The 10 Sox2 interaction clusters. Cytoscape diagram of the Sox2 interacting protein clusters differentially coloured based on the reported functions of these interactors using STRING.

A multitude of proteins responsible for performing various types of post-translational modifications were identified as part of the Sox2 protein interaction network. Listed in Table 3.2 are the 7 kinases, 2 cullins, 1 E3 ubiquitin ligase, 2 E2 ubiquitin-conjugating enzymes and 2 ubiquitin peptidases identified by IP-LC-MS/MS.

Protein Class	Gene Name	Number of Peptides Detected	% Sequence Coverage	Intensity Ratio (sample/control)
Kinase	ATM	2	1	*
	Map3k (ERK1)	3	9	*
	BUB1	4	5	28
	Melk	17	31	23
	Cit	4	2	*
	Chka	3	5	*
	Sgk223	11	10	*
Cullins	CUL3	11	17	*
	CUL4B	7	8	*
E3 Ub Ligase	Huwe1	9	3	*
E2 Ub Conjugating	Ube2o	5	4	*
	Ube2s	2	9	*
Ub Peptidase	Usp34	11	13	113
	Usp28	7	2	*

*** Indicates that a value of 0 was detected in the control**

Table 3.2: Sox2 interacting proteins involved in post-translational modifications based on n=2 experiments. Summarized are the proteins identified by IP-LC-MS/MS in this study to co-purify with Sox2 protein. The intensity ratio (sample/control) reported here is an intensity measure using protein intensity extracted by MaxQuant software which corresponds to the sum of the intensity observed for the MS/MS-identified peptides composing the protein.

3.3.3 Esrrβ protein-protein interaction network

In this study, endogenous Esrrβ protein complexes were immunopurified using a monoclonal anti-Esrrβ antibody coupled to Protein A/G agarose beads. The entire sample lanes were cut from the gel (Shown in Figure 3.9) and digested with trypsin for subsequent LC-MS/MS analysis. The results are based on 2 biological replicates of anti-Esrrβ Protein A/G Plus agarose IPs from E14tg2a whole cell lysates compared to 2 biological replicates of anti-mouse IgG agarose IPs from E14tg2a whole cell lysates.

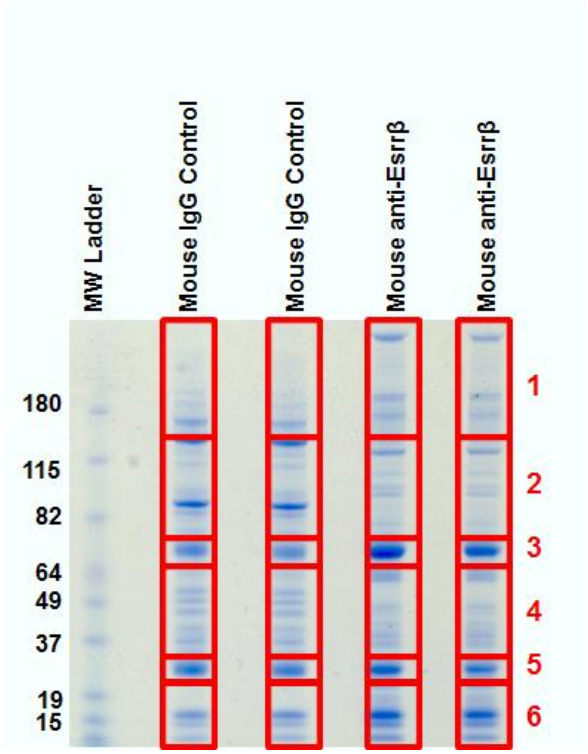


Figure 3.9: Immunopurification of endogenous Esrrβ protein complexes. Coomassie-stained gel of endogenous protein complexes isolated using anti-mouse-Esrrβ monoclonal antibody coupled to Protein A/G Plus agarose beads. Gel slices excised for digestion and subsequent LC-MS/MS analysis are indicated by red boxes and numbers.

Of the 187 high confidence interactors identified in the Esrrβ immunopurifications (see APPENDIX C), 88 were grouped into 10 protein interaction clusters based on reported interactions between these proteins. The interacting protein clusters were coloured based on their known function (reported by

STRING) and the bright green-coloured nodes represent the other 97 proteins identified to interact with Esrrβ (Figure 3.10)

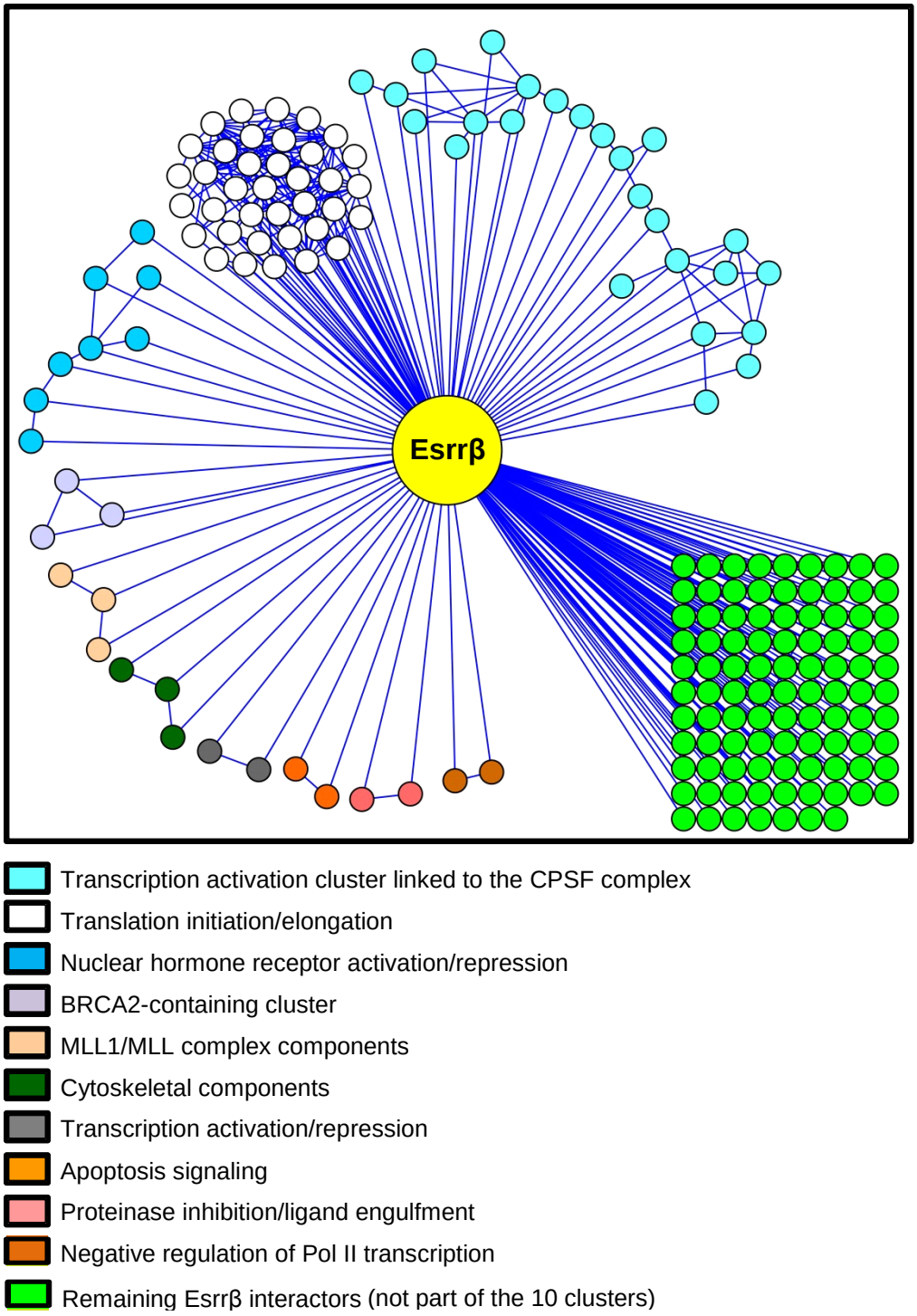


Figure 3.10: The Esrrβ protein-protein interaction network generated in this study. Cytoscape diagram of the high confidence Esrrβ interacting proteins identified by LC-MS/MS. Proteins that are reported to interact with each other are clustered together and differentially coloured based on the known function of these interacting groups as reported by STRING software. Each of the protein clusters are further illustrated in Figures 3.11 and 3.12.

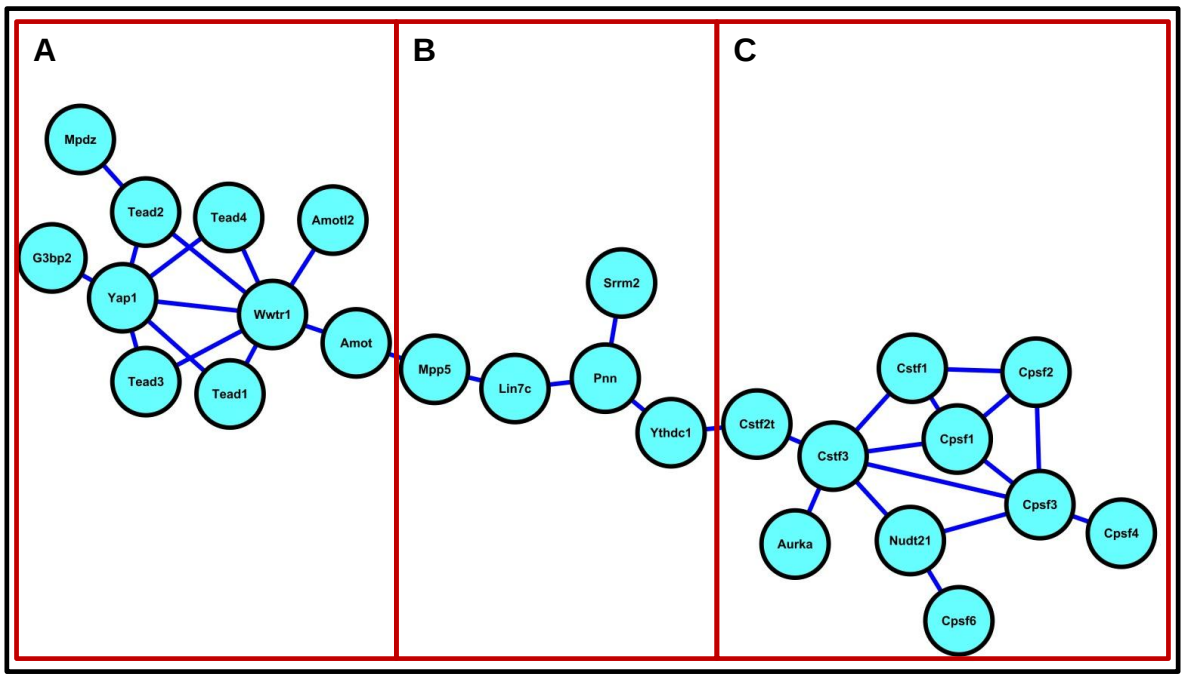
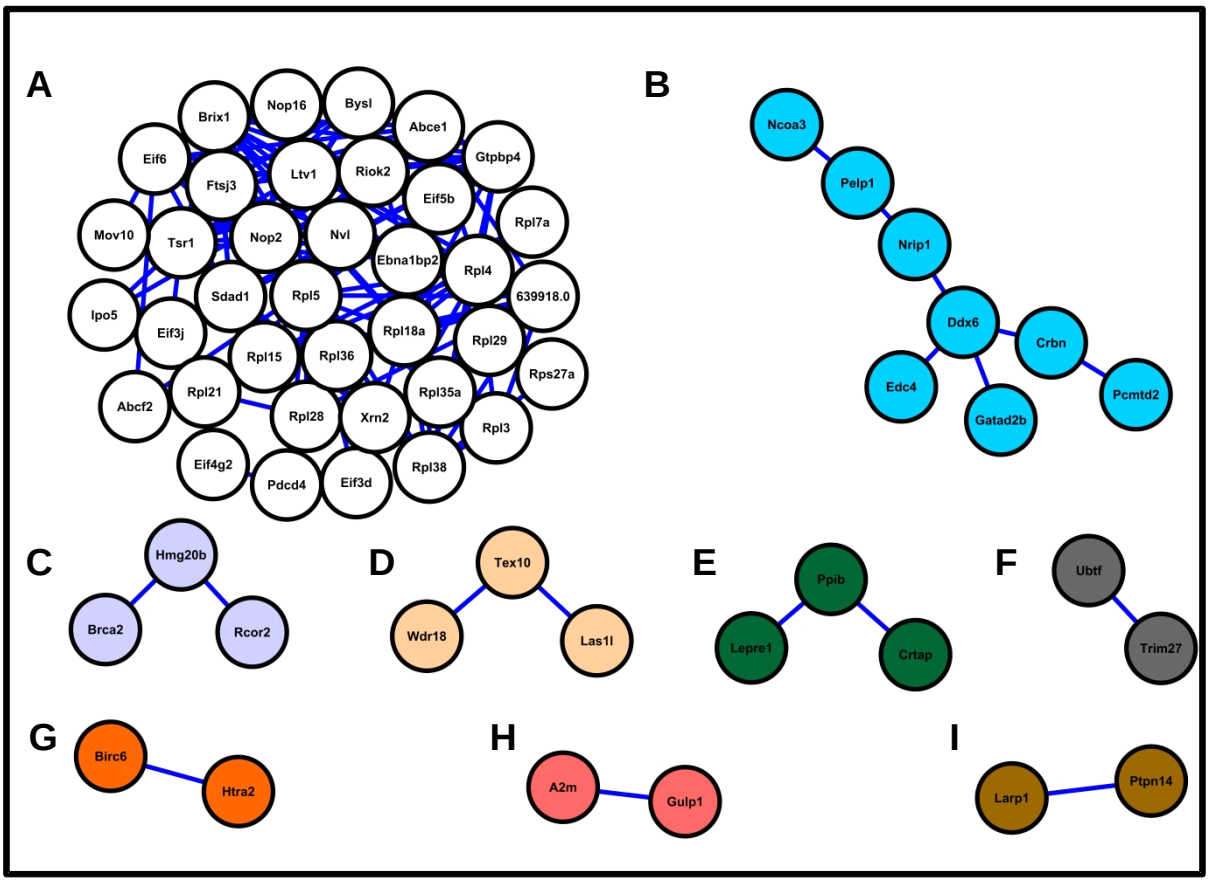


Figure 3.11: The transcription activation cluster linked to the CPSF complex. Cytoscape diagram of a large cluster of Esrrβ interacting proteins. **A)** Subgroup of proteins containing the Tead transcription activation/repression group of proteins and links to mRNA metabolism, **B)** String of proteins involved in mRNA metabolism and synthesis and, **C)** Cluster of proteins involved in mRNA splicing (The CPSF complex).

Shown in Figure 3.11 is the transcription activation cluster linked to the cleavage and polyadenylation specificity factor (CPSF) complex. This large group of proteins can be sub-divided into 3 sub-groups. Shown in Figure 3.11 (A) is the first sub-group which contains Tead1, Tead2, Tead3 and Tead4 which are all components of the Tead transcription activation/repression complex. (B) Shows the proteins Lin7c, Srrm2 and Pnn which are involved in mRNA metabolism. The association between Mpp5 and Amot links one half of this cluster to the Tead complex-containing cluster in (A) and the interaction between Ythcd1 and Cstf2t links this cluster to the CPSF complex-containing cluster. (C) A cluster of proteins involved in mRNA splicing (CPSF complex). Shown in this cluster are all known components of the CPSF, as well as, AurkA (Aurora kinase A).



A. mRNA translation initiation/elongation cluster.
B. Nuclear hormone receptor activation cluster linked to transcriptional repression.
C. BRCA2-Hmg20b-RCOR2 cluster.
D. MLL1/MLL complex components.
E. Cytoskeletal/membrane signaling components.
F. Transcription activation/repression.
G. Apoptosis induction/inhibition complex.
H. Proteinase inhibition cluster
I. Negative regulation of Pol II transcription complex.

Figure 3.12: The remaining 9 Esrrβ interaction clusters. Cytoscape diagram of the Esrrβ interacting protein clusters differentially coloured based on the reported functions of these interactors using STRING.

The remaining 9 interaction clusters are shown in Figure 3.12 (A-I). (A) Shown is a large cluster of ribosomal proteins involved in mRNA translation initiation and elongation. (B) Shown is a cluster containing the nuclear hormone co-activator (Ncoa3) connected to the transcriptional repressor Gatad2b. (C) Shown is the BRCA2-Hmg20b-RCOR2 cluster. (D) Shown is a cluster containing component of the MLL1/MLL complex. (E) Shown is a cluster of cytoskeletal and membrane

signaling proteins. (F) Shown are Ubtf and Trim27: a pair of proteins involved in transcription activation/repression. (G) Shown are Birc6 and Htra2 which are components of an apoptosis inhibition/induction complex. (H) Shown is a proteinase inhibition cluster of proteins. (I) Shown are the proteins Larp1 and Ptpn14 which are involved in the negative regulation of the Pol II transcription complex.

Listed in Table 3.3 are those proteins involved in performing post-translational modifications that were identified as part of the Esrrβ protein interaction network. Identified in this study were 3 kinases, 1 cullin, 1 E3 ubiquitin ligase and 1 E3 SUMO ligase.

Protein Class	Gene Name	Number of Peptides Detected	% Sequence Coverage	Intensity Ratio (sample/control)
Kinase	AurkA	11	32	*
	Skiv2l	5	5	*
	Riok2	3	6	*
Cullins	CUL7	7	4	*
E3 Ub Ligase	Ubr5	14	6	*
E3 SUMO Ligase	Ranbp2	17	7	*

*** Indicates that a value of 0 was detected in the control**

Table 3.3: Esrrβ interacting proteins involved in post-translational modifications based on n=2 experiments. Summarized are the proteins identified by IP-LC-MS/MS in this study to co-purify with Esrrβ protein. The intensity ratio (sample/control) reported here is an intensity measure using protein intensity extracted by MaxQuant software which corresponds to the sum of the intensity observed for the MS/MS-identified peptides composing the protein.

3.3.4 Sall4 protein-protein interaction network

Using a polyclonal antibody targeting endogenous Sall4 epitopes coupled to Protein A/G Plus agarose beads, a Sall4 protein interaction network was established. Shown in Figure 3.13 are the endogenous Sall4 protein complexes isolated in this study that were cut from the gel and digested with trypsin for subsequent LC-MS/MS analysis. The results are based on 2 biological replicates of anti-Sall4 Protein A/G Plus agarose IPs from E14tg2a whole cell lysates compared to 2 biological replicates of anti-goat IgG agarose IPs from E14tg2a whole cell lysates.

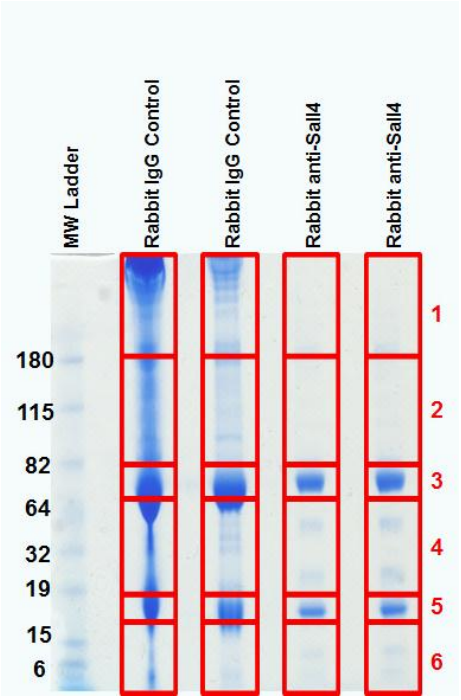


Figure 3.13: Immunopurification of endogenous Sall4 protein complexes. Coomassie-stained gel of endogenous protein complexes isolated using rabbit polyclonal antibody to mouse Sall4 coupled to Protein A/G Plus agarose beads. Gel slices excised for digestion and subsequent LC-MS/MS analysis are indicated by red boxes and numbers.

The Sall4 protein-protein interaction network identified in these experiments is composed of 104 nodes and 169 edges which are illustrated in Figure 3.14. Of the 104 high confidence interactors identified in the Sall4 immunopurifications (see APPENDIX D), 52 were grouped into 7 protein interaction clusters based on

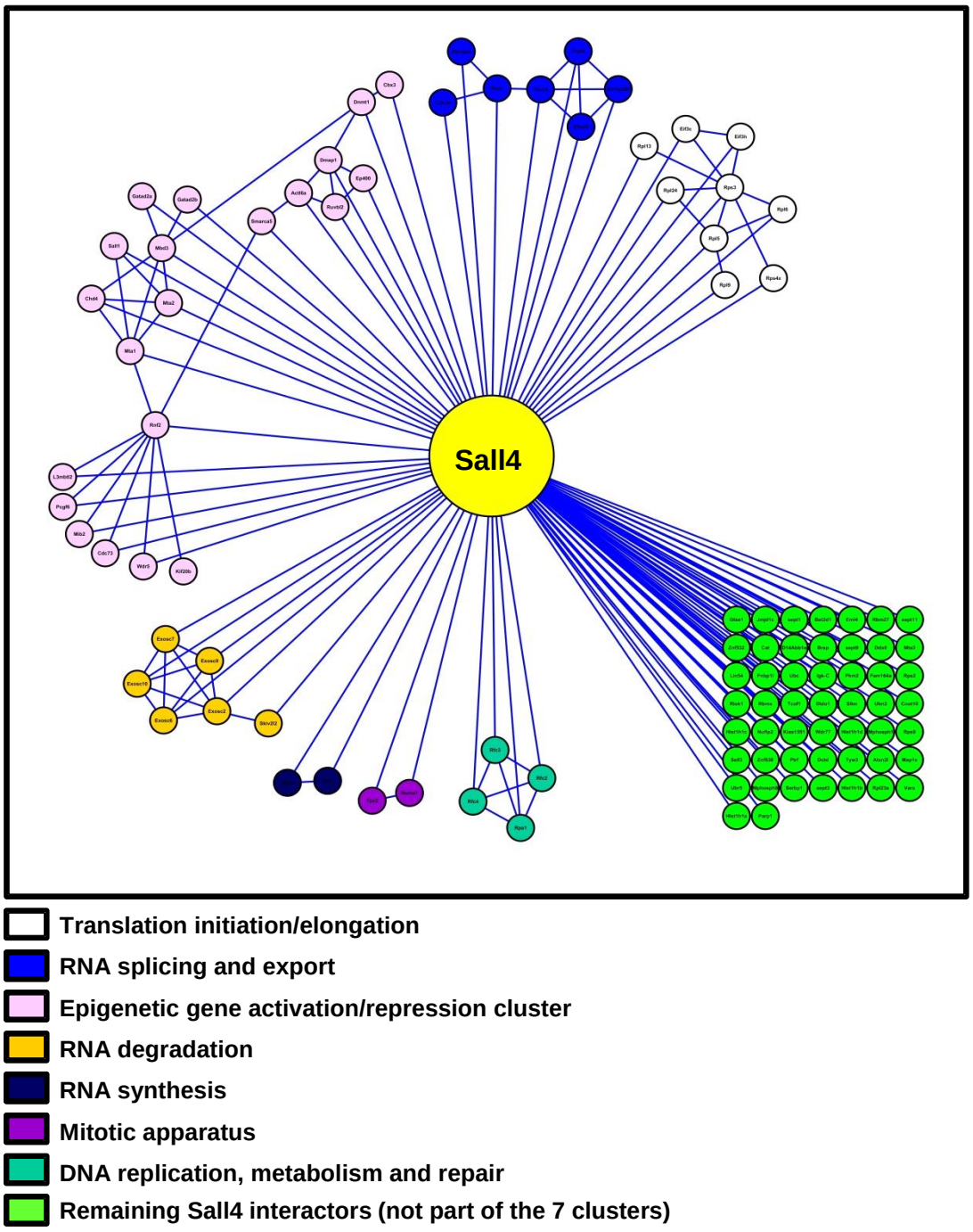
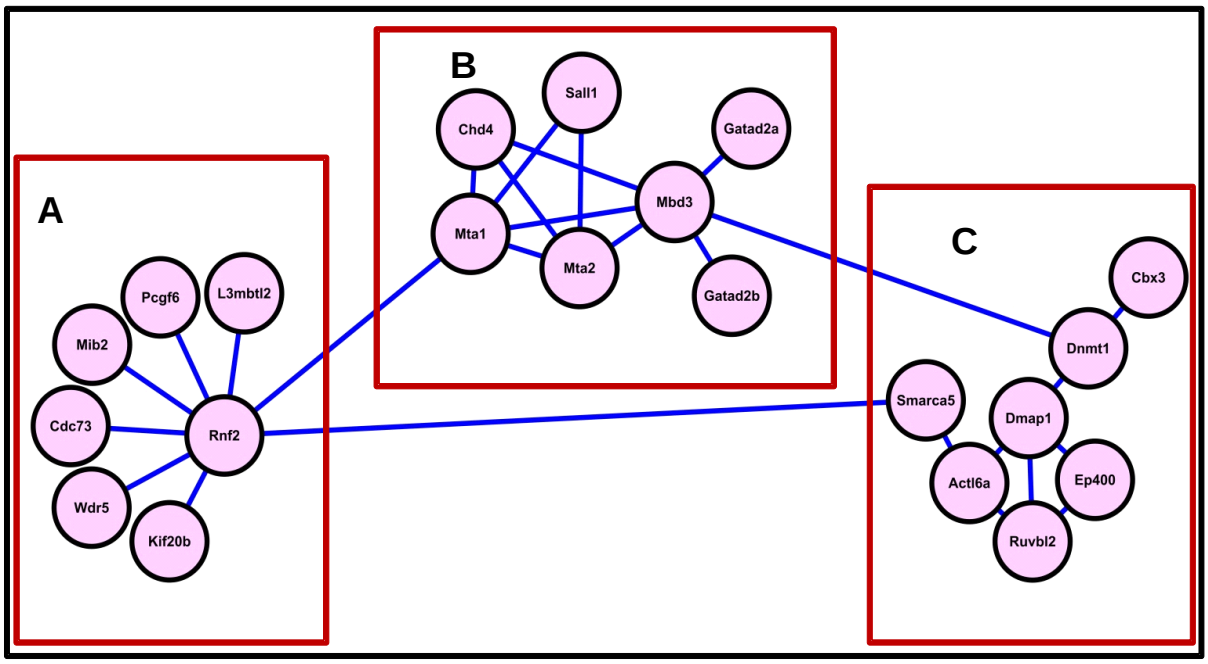


Figure 3.14: The Sall4 protein-protein interaction network generated in this study. Cytoscape diagram of the high confidence Sall4 interacting proteins identified by LC-MS/MS. Proteins that are reported to interact with each other are clustered together and differentially coloured based on the known function of these interacting groups as reported by STRING. Each of the protein clusters are further illustrated in Figures 3.15 and 3.16.

protein clusters were coloured based on their known function and the bright green-coloured nodes represent the other 51 proteins identified here as interactors with Sall4.

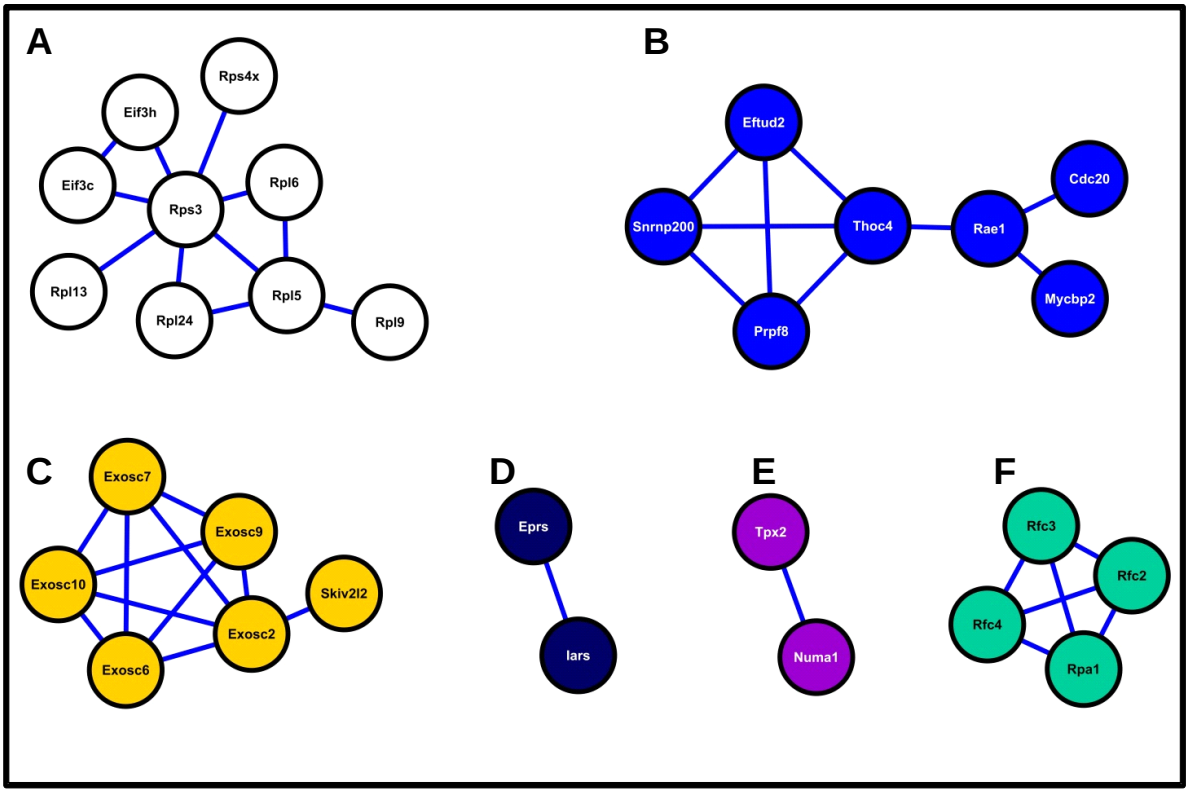


A. E3 ubiquitin-protein ligase-containing cluster involved in the mono-ubiquitination of Lys-119 on H2A.
B. DNA binding cluster involved in the recruitment of epigenetic modifying proteins to specific loci in the
C. DNA methylation and histone acetylation cluster.

Figure 3.15: The epigenetic gene activation/repression cluster. A) Epigenetic transcriptional repression and X-chromosome inactivation cluster, **B)** Cluster of DNA binding proteins bound with Mbd3 and, **C)** Subgroup containing enzymes responsible for transcriptional repression in connection with chromatin remodeling.

Shown in Figure 3.15 is the large epigenetic gene activation/repression cluster. This large group of proteins can be further sub-divided into 3 groups. Figure 3.15(A) Shows a sub-group containing the E3 ubiquitin-protein ligase Rnf2 which is involved in the mono-ubiquitination on Lys119 of Histone H2A. This protein acts in concert with the other components of this cluster (Wdr5, Cdc73 and Pcgf6) to epigenetically silence gene transcription. (B) Shows a cluster containing the majority of the NuRD complex members (Gatad2a, Gatad2b, Mbd3, Mta1 and Mta2) involved in the recruitment of epigenetic modifying enzymes to specific loci in the nucleus and, (C) Shows a cluster of proteins containing enzymes involved in epigenetic

modifications. Enzymes involved in both DNA methylation (Dnmt1) and histone acetylation (Ep400) are identified here.



A. Translation initiation/elongation cluster.
B. mRNA splicing and export cluster containing an essential APC E3 ligase
C. Exosome complex involved in mRNA degradation.
D. tRNA synthesis cluster.
E. Nuclear mitotic apparatus cluster.
F. Replication factor C

Figure 3.16: The remaining 6 Sall4 interaction clusters. Cytoscape diagram of the Sall4 interacting protein clusters differentially coloured based on the reported functions of these interactors using STRING.

The remaining 9 interaction clusters are shown in Figure 3.16 (A-F). (A) Shown is a cluster of proteins involved in translation initiation and elongation. (B) Shown is an mRNA splicing and export cluster of proteins containing the essential Anaphase Promoting Complex (APC) E3 ligase component Cdc20. (C) Shown is a cluster of proteins containing the entire exosome complex which is involved in mRNA degradation. (D) Shown are Eprs and Iars proteins which are involved in tRNA synthesis. (E) Shown is a pair of proteins which are components of the nuclear

mitotic apparatus: Numa1 and Tpx2. (F) Shown is a cluster of proteins containing Rfc2, Rfc3, Rfc4 and Rpa1. Together, these proteins form the replication factor C complex.

In Table 3.4 the proteins involved in performing post-translational modifications that were identified as part of the Sall4 protein interaction network are listed. Identified in this study were 1 kinase, 4 E3 ubiquitin ligases and 1 E3 ubiquitin ligase substrate recognition component.

Protein Class	Gene Name	Number of Peptides Detected	% Sequence Coverage	Intensity Ratio (sample/control)
Kinase	Riok1	5	10	*
E3 Ub Ligase	Mib2	3	3	*
	Ubr5	7	2	*
	Rnf2	4	16	*
	MycBP2	12	3	*
	BRAP	21	38	*
Substrate Recognition	Cdc20	5	13	*

*** Indicates that a value of 0 was detected in the control**

Table 3.4: Sall4 interacting proteins involved in post-translational modifications based on n=2 experiments. Summarized are the proteins identified by IP-LC-MS/MS in this study to co-purify with Sall4 protein. The intensity ratio (sample/control) reported here is an intensity measure using protein intensity extracted by MaxQuant software which corresponds to the sum of the intensity observed for the MS/MS-identified peptides composing the protein.

3.3.5 Pi3kcd protein-protein interaction network

In this set of experiments, the bait protein used was the kinase Pik3cd with six repeats of the HA epitope tag on its amino terminus (Takahashi *et al.*, 2003). A stable stem cell line was generated which constitutively expresses 6XHA-Pi3kcd in an E14tg2a wild-type background. Figure 3.17 shows the 6XHA-Pik3cd-containing protein complexes isolated using anti-HA High Affinity matrix. Once again, because the 6XHA-Pik3cd protein is expressed from the pPyCAGIP plasmid, lysates

prepared from the TNG cell line also cultured in the presence of puromycin were used as a control cell line for these experiments. The results are based on 2 biological replicates of anti-HA High Affinity matrix IPs from 6XHA-Pik3cd whole cell lysates compared to 2 biological replicates of anti-HA High Affinity IPs from TNG whole cell lysates.

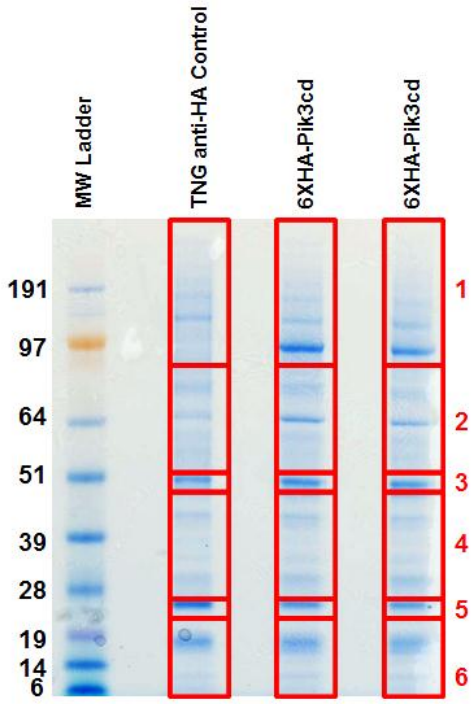


Figure 3.17: Immunopurification of Pik3cd protein complexes. Coomassie-stained gel of Pik3cd protein complexes isolated using anti-HA High Affinity matrix. Gel slices excised for digestion and subsequent LC-MS/MS analysis are indicated by red boxes and numbers.

The Pik3cd interaction network consists of 95 nodes and 387 edges which are illustrated in Figure 3.18. Of the 95 high confidence interactors identified in the 6XHA-Pik3cd immunopurifications (APPENDIX E), 44 were grouped into 5 protein interaction clusters based on the reported interactions between these proteins using STRING software (Figure 3.19). The interacting protein clusters were coloured based on their known function and the bright green-coloured nodes represent the other 51 proteins identified here as interactors with Pik3cd.

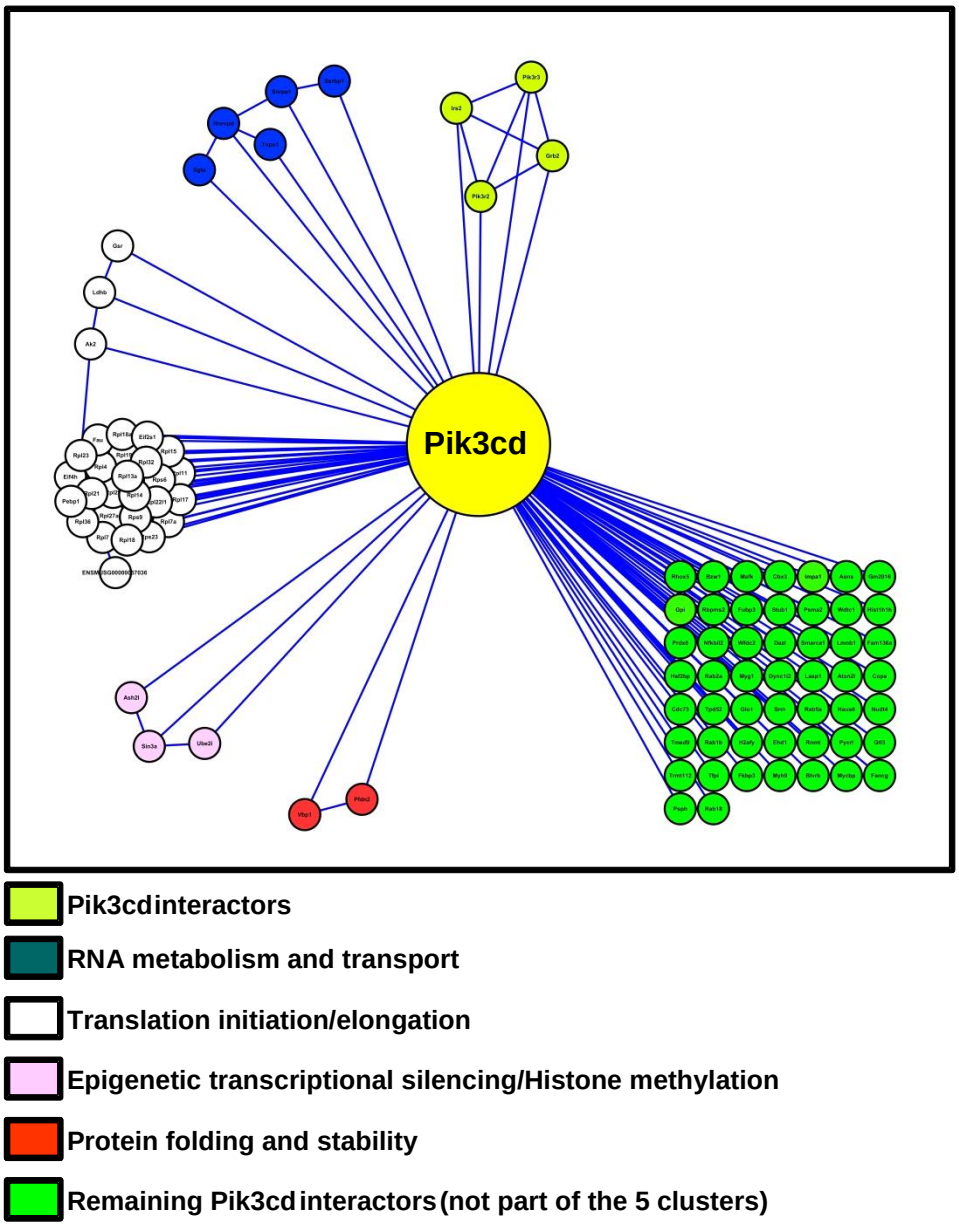
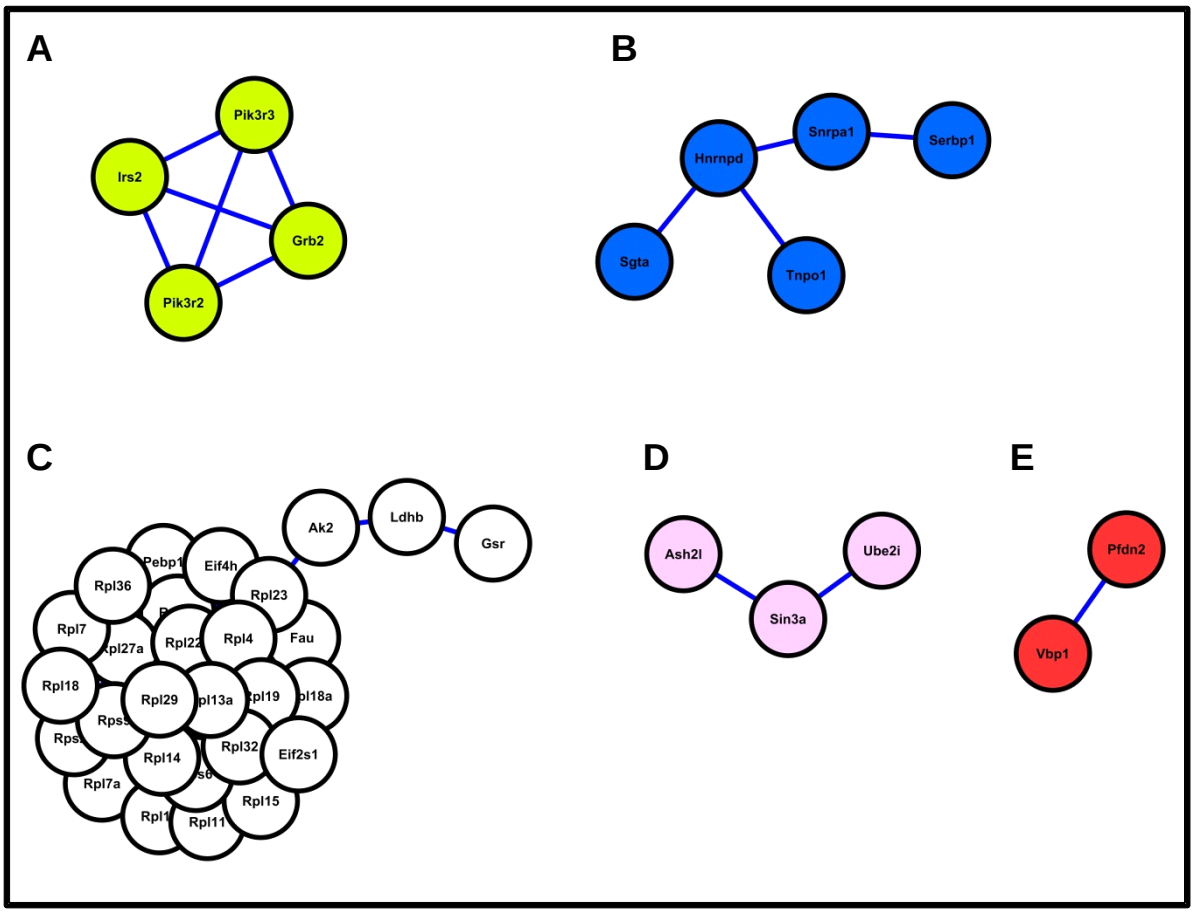


Figure 3.18: The Pik3cd protein-protein interaction network generated in this study. Cytoscape diagram of the high confidence Pik3cd interacting proteins identified by LC-MS/MS. Proteins that are reported to interact with each other are clustered together and differentially coloured based on the known function of these interacting groups as reported by STRING. Each of the protein clusters are further illustrated in Figure 3.19.



A. Pik3cd cluster containing Pik3cd regulatory units and previously identified interactors.
B. RNA metabolism and transport cluster.
C. Translation initiation/elongation cluster.
D. Epigenetic transcriptional silencing and histone methylation cluster.
E. Protein folding and stability cluster.

Figure 3.19: The 5 Pik3cd interaction clusters. Cytoscape diagram of the Pik3cd interacting protein clusters differentially coloured based on the reported functions of these interactors using STRING.

All 5 interaction clusters are shown in Figure 3.19 (A-E). (A) shown is a cluster of proteins containing the Pik3cd regulatory subunits Pik3r2, Pik3r3 as well as, Grb2 and Irs2. (B) Shown is a cluster of proteins involved in RNA metabolism and transport. (C) Shown is a large cluster of ribosomal proteins responsible for translation initiation and elongation. (D) Shown are the proteins mSin3a, Ash2l and Ube2i which are known for their involvement in epigenetic transcriptional silencing and histone methylation. (E) Shown are Vpb1 and Pfdn2 which are involved in protein folding and the targeting of proteins to cytosolic Chaperonin.

Proteins involved in performing post-translational modifications that were identified as part of the Pik3cd protein interaction network are listed in Table 3.5 Identified in this study were 3 kinases, 1 E3 ubiquitin ligase and 1 E2 ubiquitin conjugating enzyme.

Protein Class	Gene Name	Number of Peptides Detected	% Sequence Coverage	Intensity Ratio (sample/control)
Kinase	AK2	3	14	*
	Pik3r3	11	25	*
	Pik3r2	14	21	*
E3 Ub Ligase	Stub1	5	21	*
E2 Ub Ligase	Ube2i	2	5	*

*** Indicates that a value of 0 was detected in the control**

Table 3.5 Pik3cd interacting proteins involved in post-translational modifications based on n=2 experiments. Summarized are the proteins identified by IP-LC-MS/MS in this study to co-purify with Pik3cd protein. The intensity ratio (sample/control) reported here is an intensity measure using protein intensity extracted by MaxQuant software which corresponds to the sum of the intensity observed for the MS/MS-identified peptides composing the protein.

3.3.6 Integration of the interaction networks

All 5 protein interaction networks were merged using Cytoscape in order to highlight shared interactors between the networks within a larger integrated network (Figure 3.20). This integrated network consists of 614 proteins which have been detected in high confidence (Materials and Methods, Section 2.2.3.5). After analysis using STRING, 1374 protein-protein interactions were found to occur within the integrated networks.

In total, 34 proteins providing links between the Nanog protein network and one other bait protein's network were identified in this study. An additional 35 proteins provide connections between one other bait protein network and another bait protein network. These connections are illustrated in Figures 3.21 through 3.29.

The interactors shared between the Nanog and Esrrβ protein interaction networks are shown in Figures 3.21 and 3.22. The 9 proteins identified in this study as being part of both protein networks are shown in Figure 3.21. A cluster of interacting proteins, components of which are found in both networks is illustrated in Figure 3.22. The light blue-coloured nodes in this figures shows those proteins that were identified only in the Nanog network. In total, 13 proteins link the Nanog and Esrrβ networks to each other.

Figure 3.23 shows the 5 proteins found in both the Nanog and Sox2 protein interaction networks.

In Figure 3.24, the 5 proteins connecting the Nanog and Sall4 protein interaction networks together are illustrated. (A) Shows the 2 individual proteins found in both networks, (B) Shows the interactors shared by Nanog and Sall4 that cluster together to form part of a complex. The components of this cluster found in

both networks are coloured green while those specifically part of the Nanog or Sall4 interaction networks are coloured blue or purple, respectively.

Figure 3.25 illustrates the individual proteins found in both the Nanog and Pik3cd protein interaction networks.

In this study, the Nanog, Esrrβ and Sox2 protein interaction networks were found to converge on a novel protein complex shown in Figure 3.26. The red arrows in this figure indicate the direction of the bait protein and the blue lines indicate with which bait protein the shared interactors (in green) were also identified. The pink coloured node in this complex indicates that this particular protein was identified only in Esrrβ immunopurifications.

Importantly, individual proteins linking 3 networks together were also identified in this study. Figure 3.27 shows that the Cdc73 protein was identified in Nanog, Pik3cd and Sall4 protein interaction networks. The Cdc73 node is coloured orange indicating a link between the Nanog network and the two other networks.

In Figure 3.28, shows that the nuclear mitotic apparatus protein, Numa1 was identified in the Sox2, Esrrβ and Sall4 immunopurifications. The Numa1 node coloured bright blue in this figure reveals the link between these three pluripotency factors.

A large central cluster containing 62 proteins having a large degree of connectivity to all of the 5 networks elucidated in these studies is shown in Figure 3.29. This cluster is composed mainly of ribosomal proteins involved in mRNA translation initiation and elongation. Additionally, this cluster contains proteins involved in mRNA splicing and metabolism, ribosome biogenesis, nonsense-mediated mRNA decay and telomere maintenance.

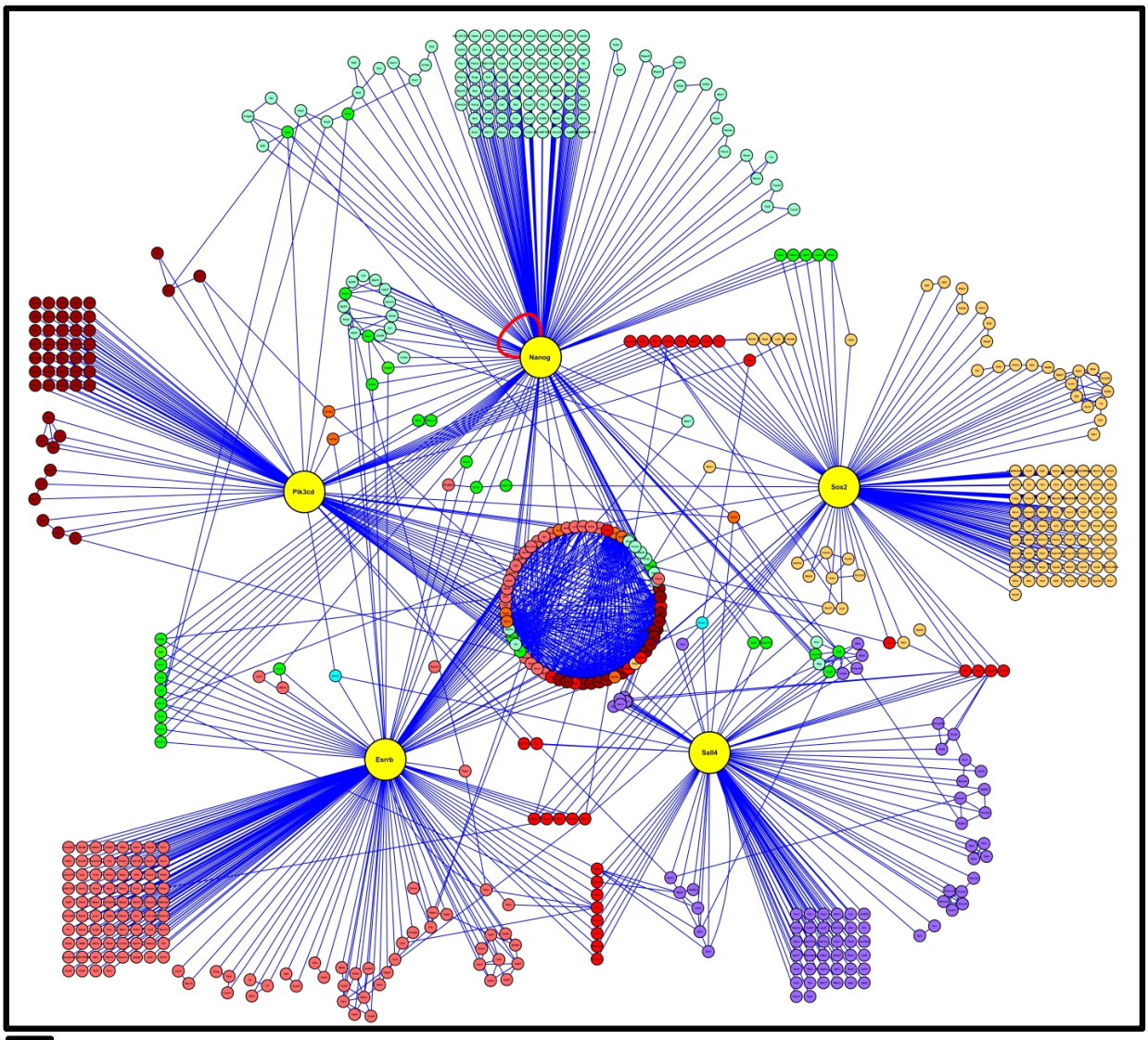


Figure 3.20: The integration of all 5 protein-protein interaction networks generated in this study. Cytoscape diagram connecting all of the protein interaction networks identified by IP-LC-MS/MS in this study. Bait proteins are coloured yellow and the nodes for each of the baits are differentially coloured. The figure legend indicates the colour scheme used to illustrate the individual networks as well as those nodes that provide links between the networks.

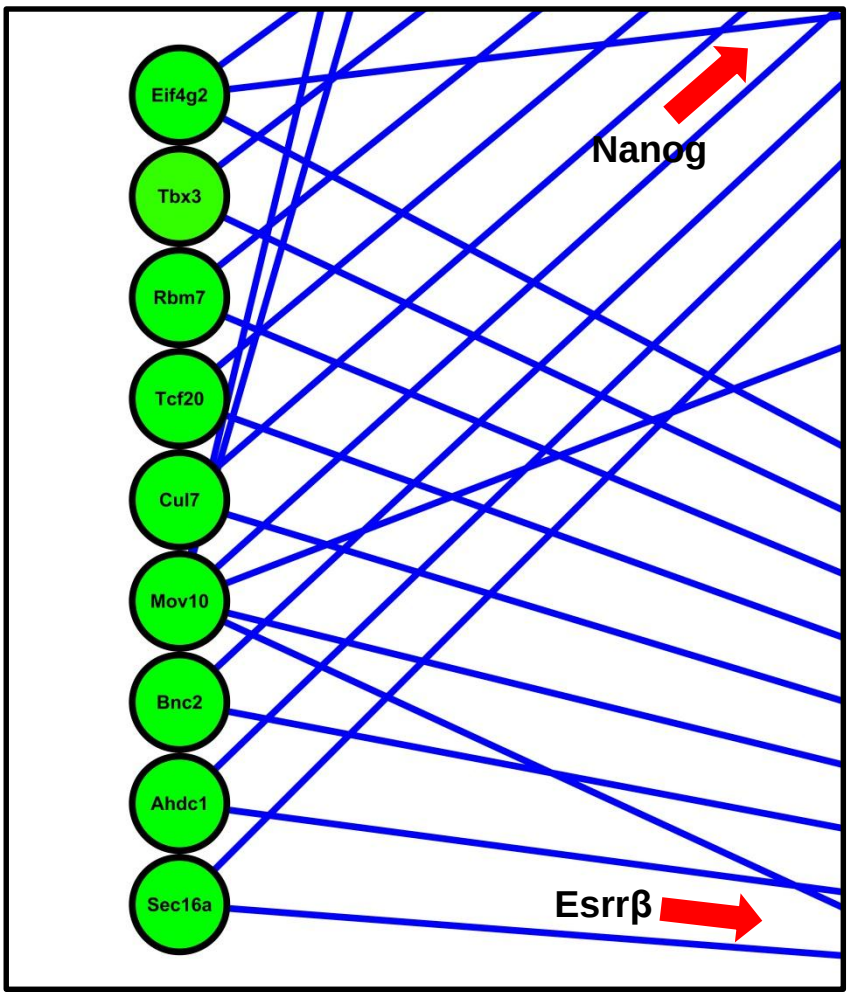


Figure 3.21: Direct connections between the Nanog and Esrrβ protein-protein interaction networks. A section of the network (Figure 3.20) illustrating a direct link between the Nanog and the Esrrβ protein-protein interaction networks (bright green). Red arrows indicate the direction of the bait protein nodes.

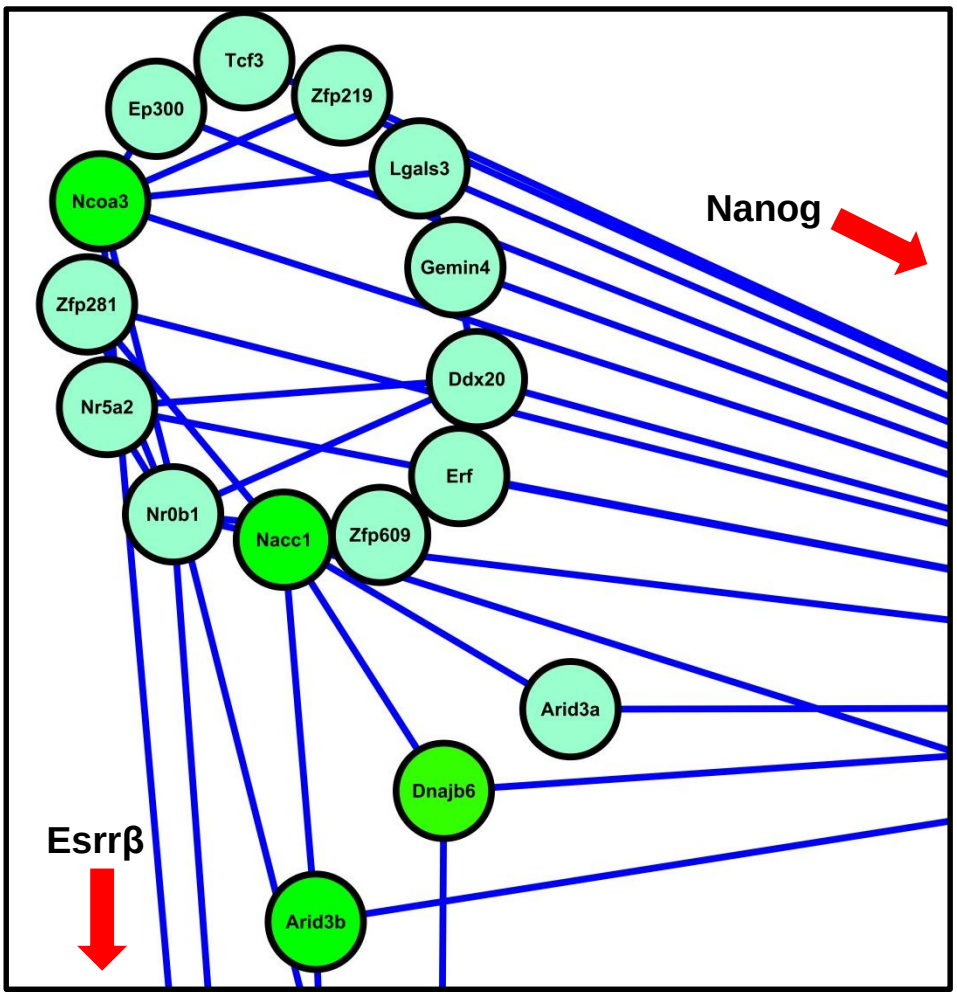


Figure 3.22: Connectivity between the Nanog and Esrrβ protein-protein interaction networks. A section of the network (Figure 3.20) illustrating a cluster of proteins linking the Nanog and the Esrrβ protein-protein interaction networks. Components of the cluster identified as interactors in both the Nanog and Esrrβ networks are shown in bright green. Red arrows indicate the direction of the bait protein nodes.

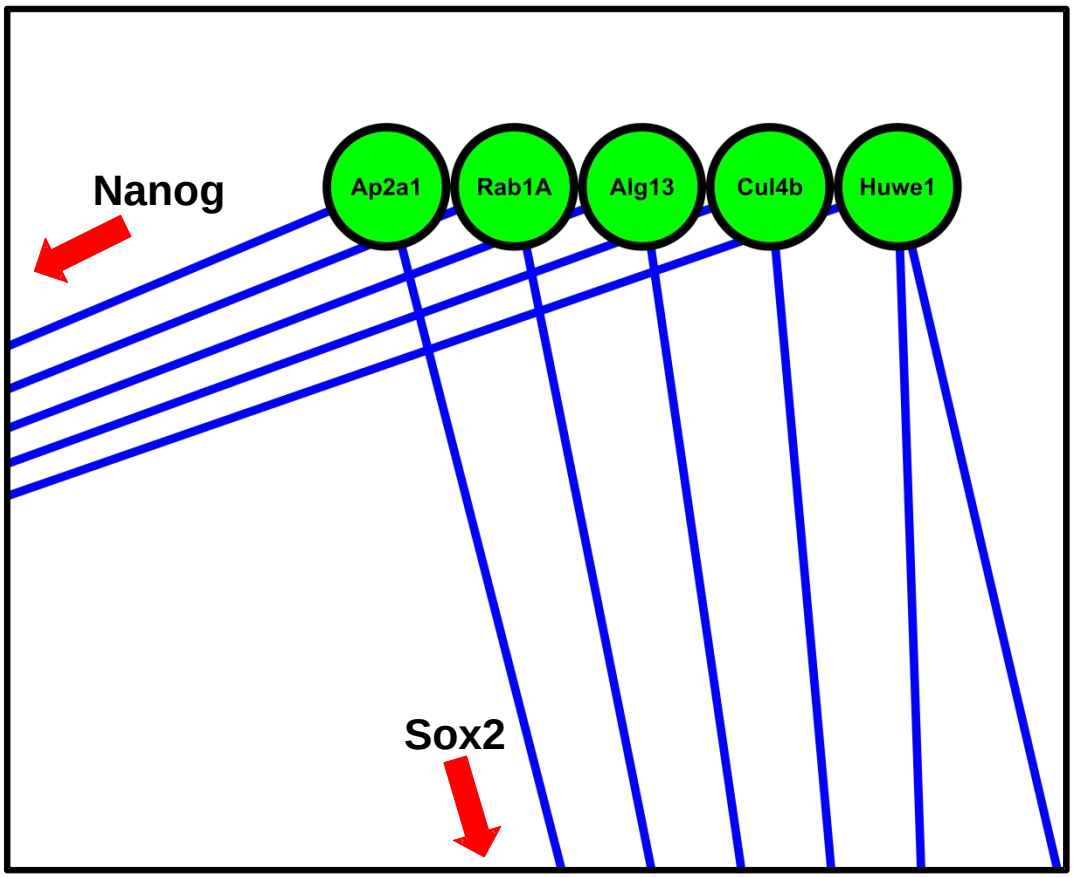


Figure 3.23: Connectivity between the Nanog and Sox2 protein-protein interaction networks. A section of the network (Figure 3.20) illustrating the proteins identified by IP-LC-MS/MS as being part of both the Nanog and the Sox2 protein-protein interaction networks (bright green). Red arrows indicate the direction of the bait protein nodes.

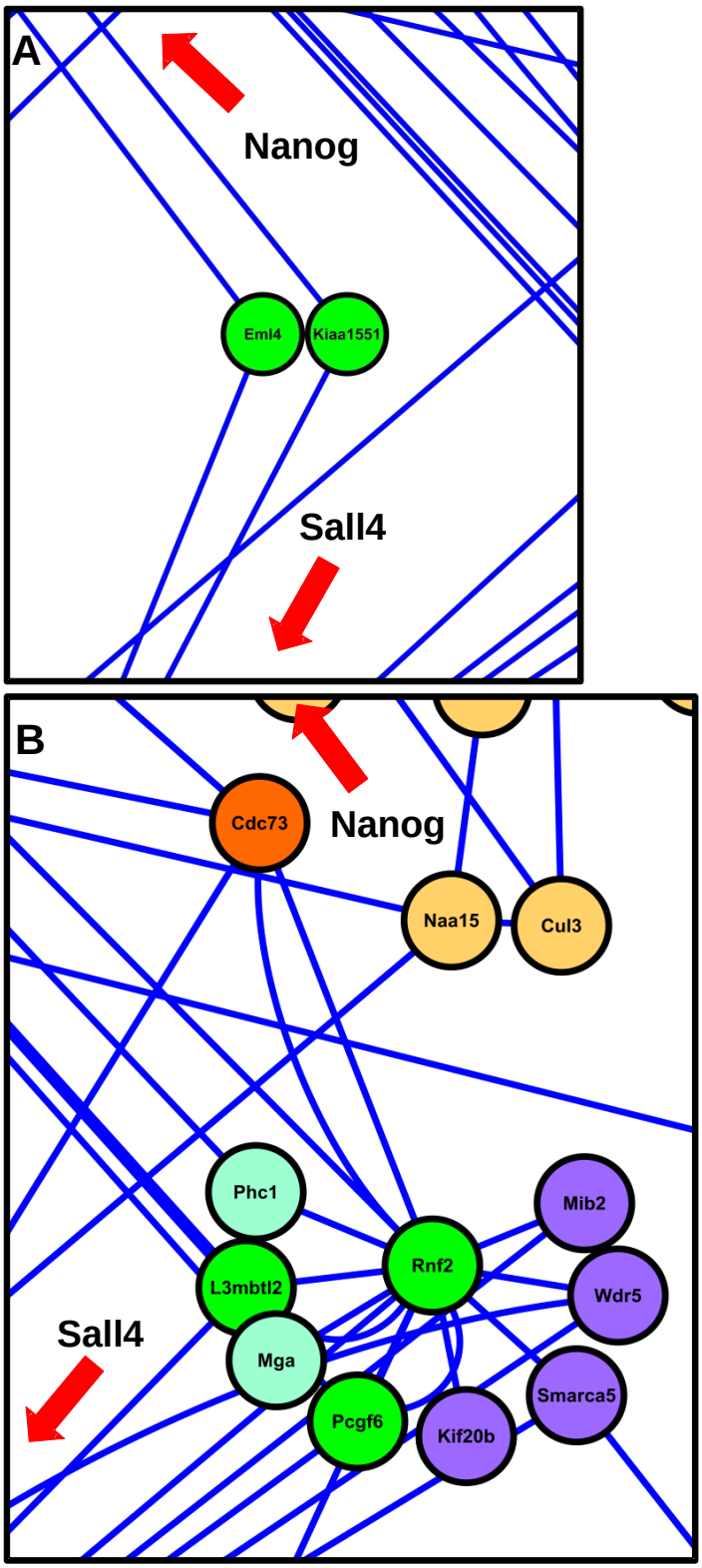


Figure 3.24: Connectivity between the Nanog and Sall4 protein-protein interaction networks. A section of the network (Figure 3.20) illustrating **(A)** the proteins identified in both the Nanog and the Sall4 protein-protein interaction networks by IP-LC-MS/MS (bright green). **(B)** Shown is a cluster of proteins in which components were identified as interacting with Nanog (light blue), Sall4 (purple) or both (bright green). Red arrows indicate the direction of the bait protein nodes.

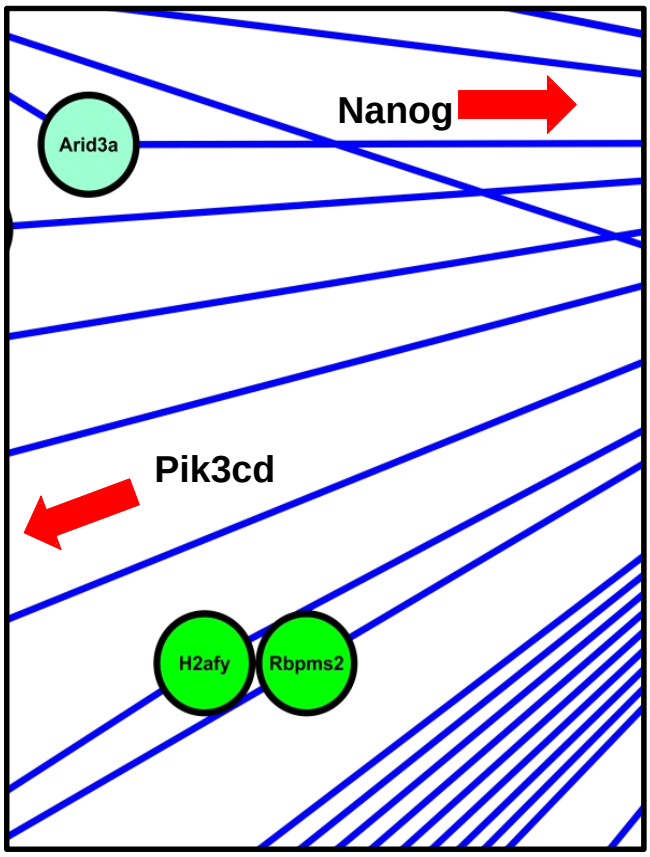


Figure 3.25: Connectivity between the Nanog and Pik3cd protein-protein interaction networks. A section of the network (Figure 3.20) illustrating the proteins identified by IP-LC-MS/MS as being part of both the Nanog and Pik3cd protein-protein interaction networks (bright green). Red arrows indicate the direction of the bait protein nodes.

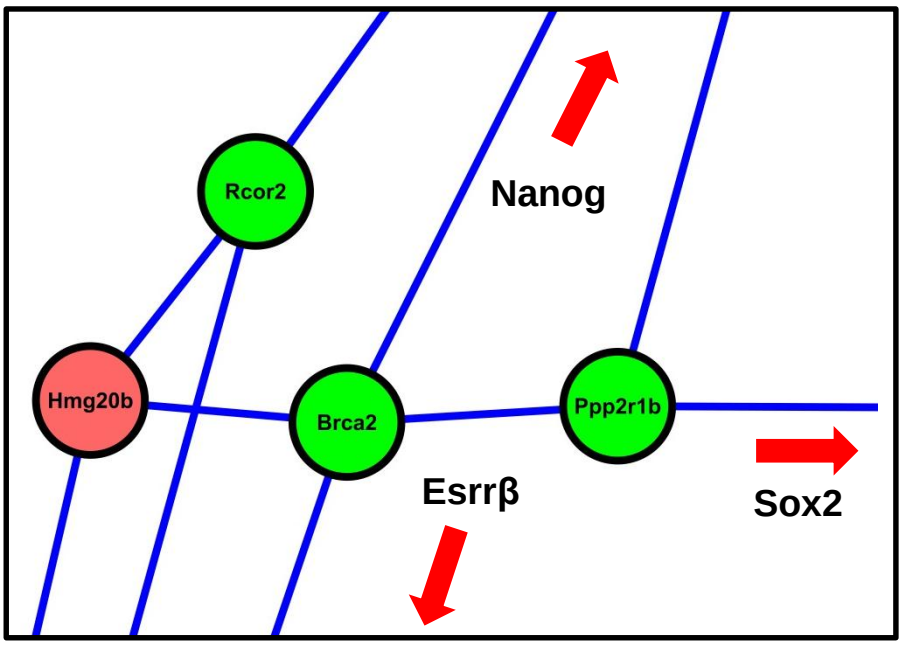


Figure 3.26: Connectivity between the Nanog, Esrrβ and Sox2 protein-protein interaction networks. A section of the network (Figure 3.20) illustrating a cluster of proteins linking the Nanog, Esrrβ and Sox2 protein-protein interaction networks. Components of the cluster providing direct links between Nanog and the two other bait proteins are shown in bright green. The pink coloured node indicates that this protein was identified only as an interactor with Esrrβ. Red arrows indicate the direction of the bait protein nodes.

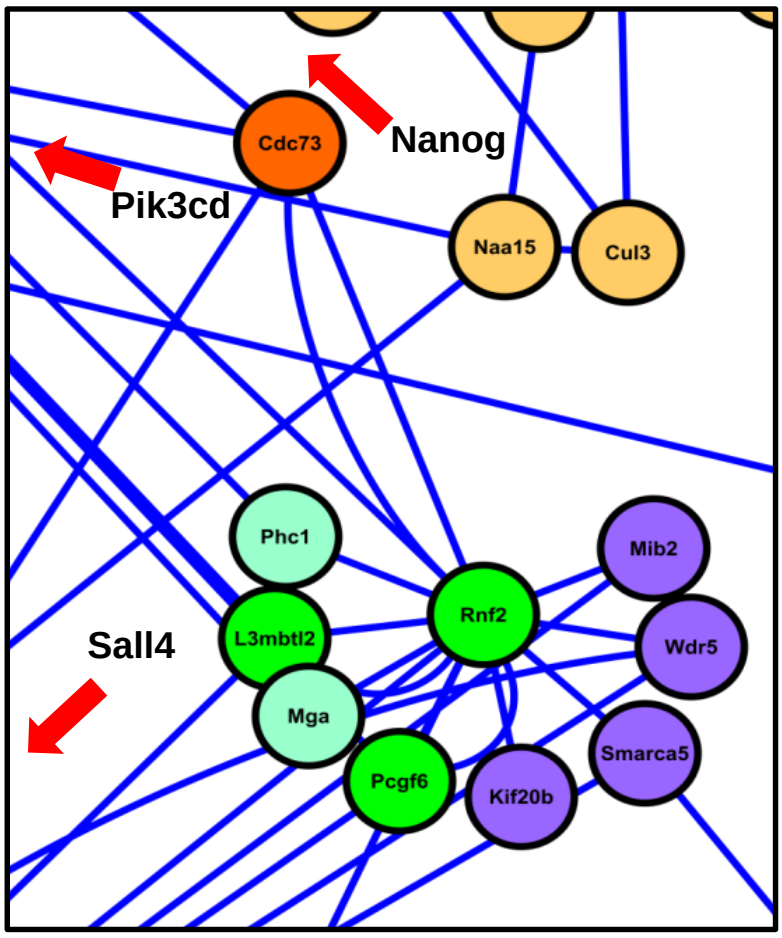


Figure 3.27: Connectivity between the Nanog, Sall4 and Pik3cd protein-protein interaction networks. A section of the network (Figure 3.20) illustrating the protein identified in the Nanog, Sall4 and Pik3cd protein-protein interaction networks by IP-LC-MS/MS (bright orange). Proteins providing links between Nanog and Sall4 are coloured bright green. Red arrows indicate the direction of the bait protein nodes.

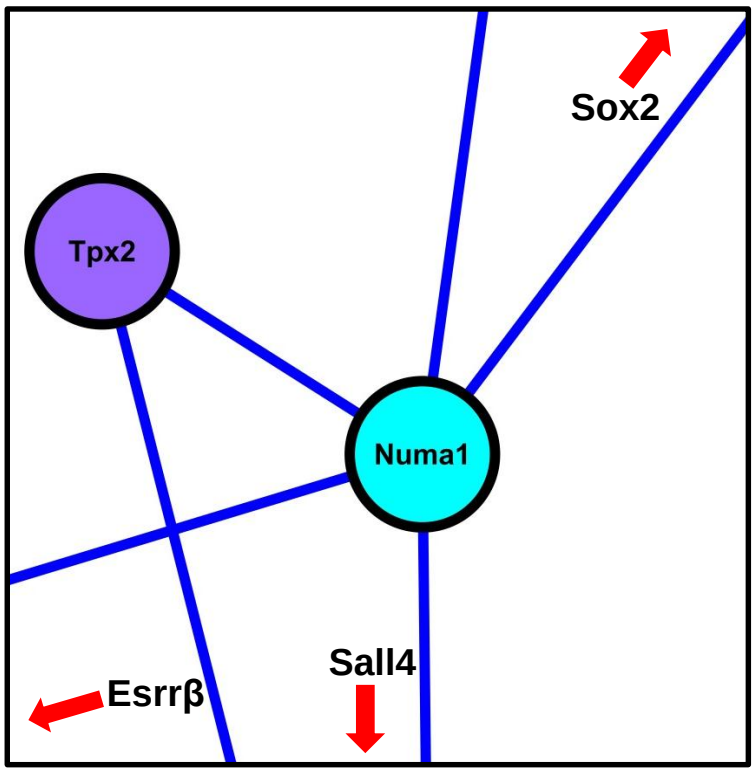


Figure 3.28: Connectivity between the Esrrβ, Sall4 and Sox2 protein-protein interaction networks. A section of the network (Figure 3.20) illustrating the protein identified by IP-LC-MS/MS in the Esrrβ, Sall4 and Sox2 protein-protein interaction networks (bright blue). Red arrows indicate the direction of the bait protein nodes.

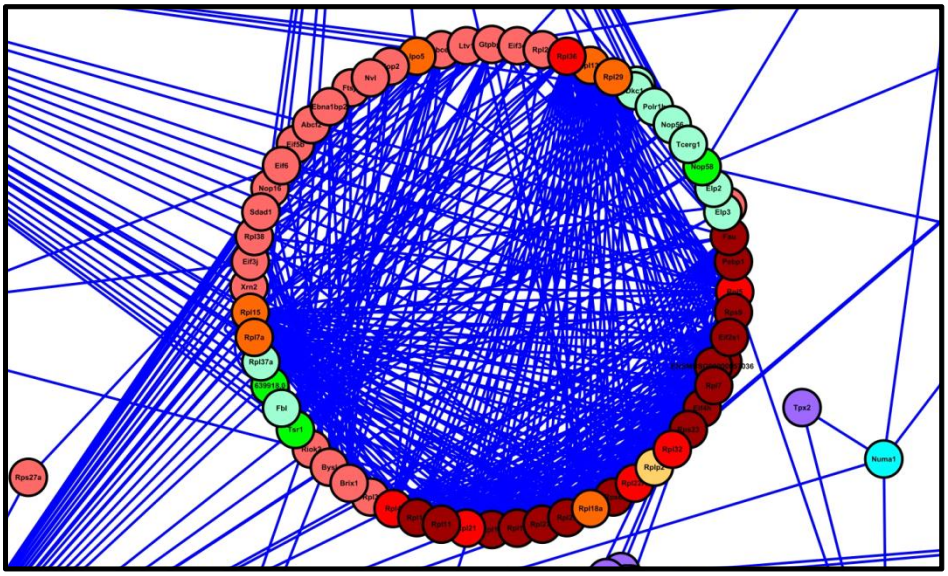


Figure 3.29: Connectivity between all 5 protein-protein interaction networks generated in this study. A section of the network (Figure 3.20) illustrating a large cluster of 63 proteins identified in all of the protein-protein interaction networks.

3.3.7 Comparison with the previously reported interaction networks of Orkin, Van den Berg and Pardo

In Figure 3.30, a Venn diagram comparing the Nanog protein-protein interaction dataset obtained in these studies with the Nanog dataset published by Stuart Orkin's lab in 2006 is shown. The Orkin dataset reports 18 proteins interacting with the Nanog protein in contrast to the 174 reported in this study.

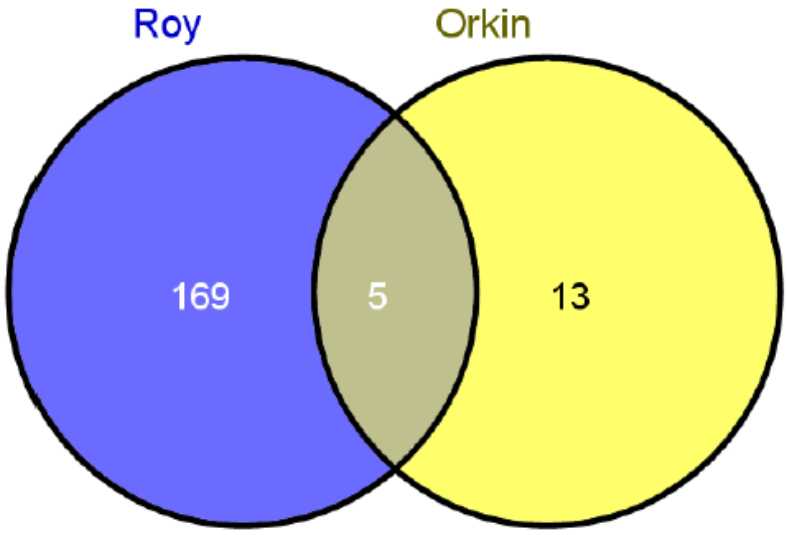


Figure 3.30: Comparison of Nanog protein interactors identified by IP-LC-MS/MS in this study with the published Orkin Nanog dataset. Venn diagram showing the overlap between the high confidence interactors with the Nanog dataset reported by Orkin.

A comparison between the published protein-protein interaction datasets obtained for Esrrβ and Sall4 in this study with the datasets previously published by Van den Berg in 2008 is shown in Figures 3.31 and 3.32. In Figure 3.31, a Venn diagram shows an overlap of only 8 proteins between the two different Esrrβ IP-LC-MS/MS datasets. A further 179 proteins are unique to the dataset acquired in this study and 74 are unique to the Van den Berg dataset. In Figure 3.32, a Venn diagram shows an overlap of 13 proteins between the two datasets obtained for Sall4. An

additional 28 proteins are unique to the Van den Berg study and 91 are unique to this study.

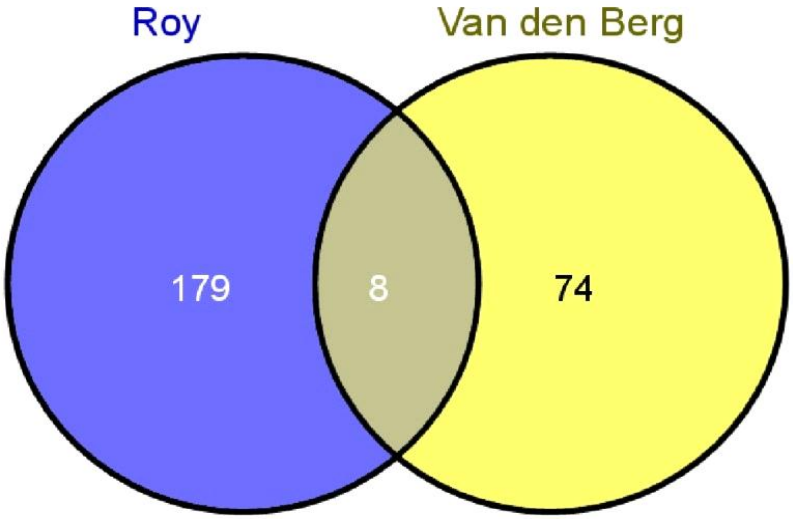


Figure 3.31: Comparison of Esrrβ protein interactors identified by IP-LC-MS/MS in this study with the published Van den Berg Esrrβ dataset. Venn diagram showing the overlap between the high confidence interactors with the Esrrβ dataset reported by Van den Berg.

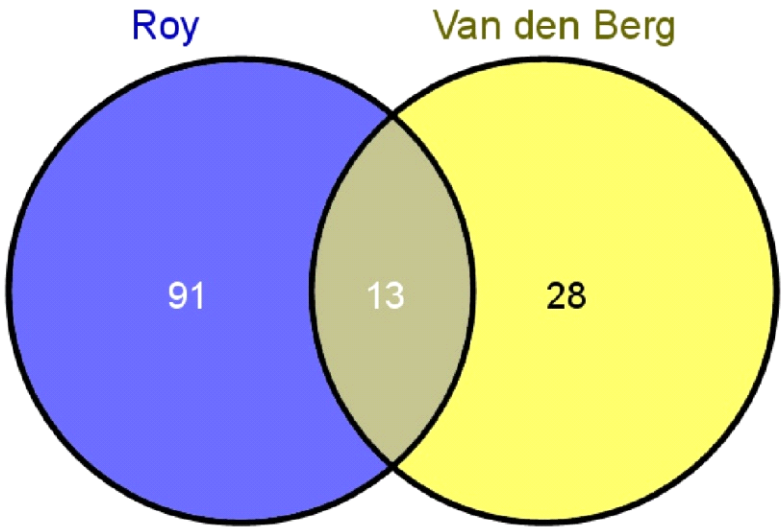
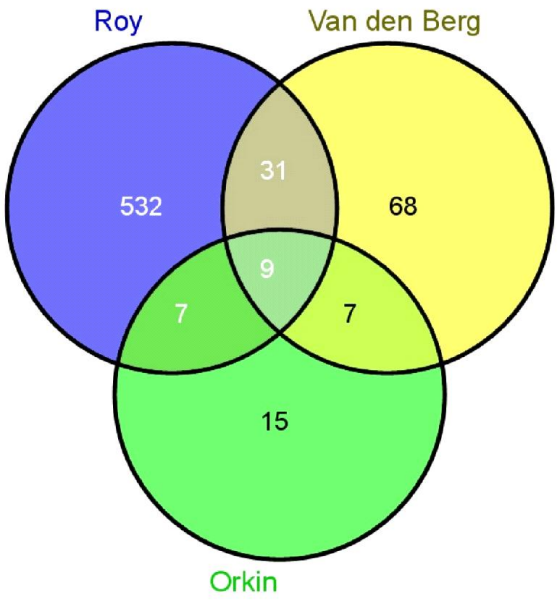


Figure 3.32: Comparison of Sall4 protein interactors identified by IP-LC-MS/MS in this study with the published Van den Berg Sall4 dataset. Venn diagram showing the overlap between the high confidence interactors with the Sall4 dataset reported by Van den Berg.

In figure 3.33, the entire IP-LC-MS/MS protein interaction datasets for all 5 baits used by Orkin, Van den Berg and in this study were compared. (A) Venn

diagram showing the overlap between all 3 networks. (B) Table summarizing the different baits used in all 3 studies.

A



B

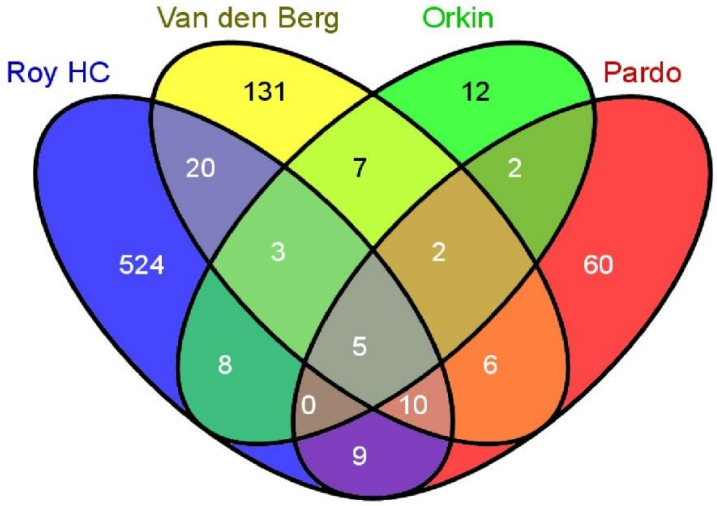
Baits	Roy (this study)	Orkin	Van den Berg
Nanog	✓	✓	
Sox2	✓		
Esrrβ	✓		✓
Sall4	✓		✓
Zfp281		✓	
Oct4		✓	✓
Nacc1		✓	
Nr0b1			✓
Tcfcp2l1			✓
Pik3cd	✓		
Zfp42		✓	

1) Wang *et al.* 2006
2) Van den Berg *et al., 2010*

Figure 3.33: Comparison of the entire protein interaction dataset from these studies with published datasets generated using the same number of baits. A) Venn diagram showing the overlap between the entire Roy High Confidence dataset with the entire datasets from Orkin and Van den Berg, **B)** Table listing the bait proteins used to obtain the Roy, Orkin and Van den Berg datasets.

The protein interactions observed in this study were obtained from whole cell lysates versus nuclear extracts (used by Orkin and Van den Berg). Therefore, it was thought that comparison with a dataset that was also obtained using whole cell lysates might provide a greater overlap of interactors. To address this question, a comparison was made with the published IP-LC-MS/MS dataset for Oct4 obtained from whole cell lysates of mESCs published by Pardo in 2010 (Pardo *et al.*, 2010).

A



B

Roy/Van den Berg/ Orkin/Pardo	Roy/Van den Berg	Roy/Orkin	Roy/Pardo
Arid3b	Akap8	Esrrb	Arid3b
Afp219	Rbm14	Nacc1	Zfp219
Gatad2b	Sox2	Rnf2	Gatad2b
Sall4	Mga	Arid3b	Sall4
Sall1	Phc1	Zfp219	Sall1
	Med12	Gatad2b	Rcor2
	Rbpj	Sall4	Mta2
	Snw1	Sall1	Chd4
	Pcgf6	Nr0b1	Smarca5
	Polr2b	Zfp281	Gatad2a
	Polr2a	Ilf2	Mta1
	Polr2c	Arid3a	Sall3
	Zfp462	Rai14	Mbd3
	Ep400	Wapal	Mta3
	Trrap	Wdr18	Actl6a
	Ruvbl2	Nanog	Cul4b
	Dmap1		Trim33
	Wdr5		Trim24
	Sin3a		Amotl2
	EsrrB		Acin1
	Nacc1		P4ha1
	Rnf2		Ubn2
	Arid3b		Rpa1
	Zfp219		Parp1
	Gatad2b		
	Sall4		
	Sall1		
	Rcor2		
	Mta2		
	Chd4		
	Smarca5		
	Gatad2a		
	Mta1		
	Sall3		
	Mbd3		
	Mta3		
	Actl6a		

Figure 3.34: Comparison of the entire protein interaction dataset from these studies with those of Orkin, van den Berg and Pardo. A) Venn diagram showing the overlap between the entire Roy High Confidence dataset with the entire dataset published by Orkin, Van den Berg and Pardo. **B)** Table listing the interactors shared by the Roy, Orkin, Van den Berg and Pardo datasets.

In Figure 3.34 (A) Venn diagram showing that the dataset acquired in this study has an overlap of 24 proteins with the Pardo dataset, 38 proteins with the Van den Berg dataset and 16 proteins with Orkin dataset. While this comparison did not account for the overall lack of overlap in the datasets, it did reveal an interesting interactor shared by this study and Pardo's dataset. The table in Figure 3.34 (B) reveals that CUL4B is an interactor that is unique to the datasets acquired using whole cell lysates.

3.4 Discussion

In this chapter, the mass spectrometry data presented were obtained using previously published immunopurification protocols and tandem MS/MS methodology (Luke-Glaser *et al*., 2007; Olma *et al*., 2009). Modifications were made to these published protocols and are further detailed in the text below. In most cases, epitope-tagged pluripotency factors expressed stably from the pPyCAGIP plasmid were immunopurified from mouse embryonic stem cells using epitope-specific antibodies. In the case of two baits (Esrrβ and Sall4), however, commercially available antibodies to endogenous epitopes were employed at the immunopurification step. The datasets presented here are unique from the previously published datasets from Van den Berg and Orkin (Van den Berg *et al.*, 2010; Wang *et al.*, 2006) in that all immunopurifications have been performed on whole cell lysates rather than nuclear extracts. Additionally, the chelating agent EDTA which is often used as a phosphatase inhibitor during immunopurifications, was omitted from all buffers. The rationale for this exclusion is that EDTA, while it inhibits phosphatases, it can also inhibit kinases. The use of whole cell lysates and the exclusion of EDTA were done in order to identify kinases which may possibly

phosphorylate the bait proteins and identify signalling pathways that may regulate the stability and function of these pluripotency factors. The omission of EDTA from lysis buffer, however, requires that lysates are pre-cleared with an inert matrix (Chen *et al.*, 2007) briefly prior to the incubation with the immunoaffinity matrix.

In order to minimize the potential contribution of the gelatin matrix proteins to the samples, cells were instead transferred to non-gelatinized Cell^{+} tissue culture plates manufactured by Sarstedt one day prior to harvest for immunopurifications. Another advantage of using a gelatin-free culture environment is that it allowed for rapid lysis of cells on the dish as opposed to cell lysis following lengthy centrifugation steps in order to remove gelatin contamination. This resulted in an improved number of peptides identified for bait and prey proteins in samples and allowed for purification of epitope-tagged baits to be performed on a scale of 10 x 15cm plates per sample and for endogenous immunopurifications to be performed on a scale of 5 x 15 cm plates per sample.

3.4.1 The Nanog Protein-Protein interaction network

Nanog is often referred to as a ‘safeguard’ for pluripotency (Chambers *et al.*, 2007) due to the fact that its expression at high levels inhibits differentiation of mESCs. Conversely, mESCs can differentiate only when Nanog expression levels are low. This defining characteristic of Nanog as a ‘gate-keeper’ to pluripotency is why it was chosen as the primary bait protein as the primary bait protein around which the protein interaction network in this study was centered.

Among the 33 proteins shown in Figure 3.4, were Dkc1 (Dyskerin), Smg5 (hEST1B) and Smg7 (hEST1C) which are all involved in nonsense-mediated mRNA

decay and telomere maintenance (Ruggero *et al.*, 2003; Ohnishi *et al.*, 2003). These three proteins further strengthen link between Figure 3.4 (A) and (B).

In Figure 3.5 part A(i) the nuclear hormone receptors and Sox2-containing cluster of proteins identified in this study is illustrated. This cluster links hormone-dependent transcriptional activation/repression with two master regulators of pluripotency: Nanog and Sox2.

Figure 3.5 part A(ii), reveals a cluster of 6 proteins involved in epigenetic silencing and gene repression through chromatin remodelling interacting with Nanog. Phc1, for instance, is a key component of the Polycomb group (PcG) multi-protein PRC1 Polycomb-group repressive complex 1 (Satijn *et al.*, 1997). The PRC1 complex acts via chromatin remodelling and modification of histones; it mediates mono-ubiquitination of histone H2A 'Lys-119', rendering chromatin heritably changed in its expressibility (Satjin and Otte, 1999). Another essential component of the PRC1 complex is ring finger protein 2 (Rnf2). Rnf2 is an E3 ubiquitin-protein ligase that mediates mono-ubiquitination of Lys-119 on histone H2A, thereby playing a central role in histone code and gene regulation. H2A 'Lys-119' ubiquitination gives a specific tag for epigenetic transcriptional repression and participates in X chromosome inactivation of female mammals (Lee *et al.*, 2001).

Figure 3.5(B) demonstrates that protein complexes purified from the 3XFLAG-Nanog mES cell line contain endogenous Esrrβ and Nr0B1/Dax1 proteins. This confirms the identification of these two nuclear hormone receptors as being part of the Nanog protein-protein interaction network obtained by LC-MS/MS in this study.

Additionally, Figure 3.5(C) reveals that 3XFLAG-Nanog protein complexes co-purify with the endogenous Sox2 protein, as well as, several components of the endogenous PKA signaling pathway. This provides further confirmation that the master regulator Sox2 and those PKA signaling proteins identified by LC-MS/MS in this study are indeed part of the Nanog protein interaction landscape in mESCs.

Unique to this study is the identification of BRCA2 as a high confidence interactor with Nanog. BRCA2 is best known for its involvement in double-strand break repair and/or homologous recombination which. To our knowledge, BRCA2 has never been reported to be associated with Nanog and was identified here with 7 peptides.

Also unique to this study was the identification of CUL4B and CUL7 cullin proteins co-purifying with Nanog. This study is the first to report the potential post-translational regulation of Nanog by Cullin-RING-ligase (CRL)-mediated ubiquitination. The involvement of these proteins in the control of Nanog protein levels is explored in Chapter 4: Exploring the Post-Translational Modification of Pluripotency Factors Using Chemical Compounds and Small Molecules.

3.4.2 The Sox2 Protein-Protein interaction network

Sox2 is a High Mobility Group domain-containing transcription factor that is essential for pluripotency (Masui *et al.*, 2007*).* Since Sox2 was identified in Nanog purifications as a high confidence interactor, reciprocal Sox2 IP-MS/MS experiments were performed.

However, Nanog was identified as a low confidence interactor at similar

levels in both sample and control immunopurifications. Similarly, Esrrβ was identified as a interactor close to but above control levels. The DNA damage repair protein BRCA2 was not identified in these experiments. However, several other DNA damage repair proteins including Rad50, ATM, MMS19, Faap100 and Brip1 (BRCA1 interacting protein) were identified as high confidence interactors in these purifications. The fact that DNA damage repair proteins are also identified in this interaction network suggests that the metabolism and repair of DNA might be an essential function of these key pluripotency factors.

Further to this end, the double-strand break (DSB) DNA damage repair proteins Rad50 and ATM kinase both co-purified with Sox2 as high confidence interactors in this study. RAD50 is a component of the MRN complex, which plays a central role in DSB DNA repair, DNA recombination, maintenance of telomere integrity and meiosis. One possible explanation for the presence of RAD50 in this network may be a requirement for its ability to bind broken DNA ends and hold them in close proximity (Dolganov *et al.*, 1996) in order for enzymes involved in DNA repair to act upon them.

Many of the proteins identified in this study to be involved in the DNA damage checkpoint within cells are also components of, or are functionally linked to, the mitotic checkpoint of the cell cycle. For example, Bub1 (Budding Inhibited by Benzimidazoles 1) which is involved in the enforcement of the mitotic cell cycle checkpoint, is associated with the kinetochore and has a role in regulating the exit from mitosis and normal mitotic timing in cells (Meraldi *et al.*, 2004).

To further support the notion that Sox2 might be involved in the regulation of mitosis in ES cells is the identification of Cyclin B1 as a significant interactor in this

dataset. In total, 6 proteins which are essential to the regulation and timing of mitosis were identified in this study. Namely, they are: Pds5a, Bub1, Melk, CyclinB1, Kntc1 and ERK1. The Sox2 protein's interaction with these proteins might be hinting at a novel role for the Sox2 protein in the control of mitotic entry and exit in ES cells in a manner which supports the maintenance pluripotency of these cells.

3.4.3 The Esrrβ Protein-Protein interaction network

Esrrβ is a nuclear hormone receptor that has been shown to bind to the Oct4 promoter activating Oct4 transcription and sustaining self-renewal (Zhang *et al.*, 2008; Van den Berg *et al.*, 2008). Since Esrrβ was identified consistently in Nanog purifications as a high confidence interactor in this study, reciprocal Esrrβ IP-LC-MS/MS experiments were performed. Although Nanog was identified as a low confidence interactor in these immunopurifications, Cytoscape analysis of the Esrrβ protein interaction network revealed 13 shared interactors with the Nanog protein interaction network. Therefore, it is possible that Nanog and Esrrβ do not directly interact with each other but rather co-purify with each other due to their associations with other proteins.

Another unique aspect this present study was the identification of the RCOR2-BRCA2-Hmg20b-Ppp2r1b cluster of proteins. Components of this cluster are identified the Nanog, Esrrβ and Sox2 protein interaction networks. The possible function of this particular complex has not been elucidated and suggests once again that a role may exist for DNA damage repair proteins in the maintenance of pluripotency.

3.4.4 The Sall4 Protein-Protein interaction network

Sall4 (Sal-like 4) is a zinc finger transcription factor which activates Oct4 expression and is required for the maintenance of pluripotency (Zhang *et al.*, 2006; Yuri *et al.*, 2009). Although not identified as a high confidence interactor in the experiments presented here, Sall4 has been previously reported to interact with Nanog (Wu *et al.*, 2006; Liang *et al.*, 2008). In an attempt to isolate and characterize a potential subset of low abundance Nanog and Sall4-containing complexes, reciprocal Sall4 IP-LC-MS/MS experiments were performed. In these experiments, Nanog, Esrrβ and Sox2 were not identified as either high or low confidence interactors. However, the Sall4 protein interaction network is highly connected to the Nanog protein interaction network by 8 proteins. These network-connecting proteins are: Pcgf6, Rnf2, L3mbtl2, Eml4, Kiaa1551, Cdc73, Serbp1 and Mga.

One of the largest clusters of interactors in the Sall4 network is the epigenetic gene activation/repression cluster shown in Figure 3.15. This cluster also contains the greatest number of shared interactors with Nanog. They are: Pcgf6, Rnf2, L3mbtl2 and Cdc73 and are all involved with epigenetic silencing (Lee *et al.*, 2001; Wang *et al.*, 2004; Akasaka *et al.*, 2002; Guo *et al.*, 2009; Rozenblatt-Rosen *et al.*, 2005) . This suggests that the Sall4 and Nanog proteins may not directly physically interact but rather that they can co-purify due to each protein's independent association with a shared protein complex within ES cells.

The second largest cluster of interactors in the Sall4 IPs is the entire Replication factor C complex RFC illustrated in Figure 3.16(F). The RFC complex is responsible for elongation of primed DNA templates and has a role in transcription,

DNA repair and telomere maintenance (Wang *et al.*, 2000). Therefore, this cluster also links the Sall4 protein interaction network to DNA damage repair in mESCs.

Unique to Sall4 purifications in this study, is the identification of proteins involved in Raf/MEK inhibition, Table 3.4. MycBP2 and its known interactor BRAP (BRCA1-associated protein). BRAP (also known as IMP and Rnf52) has been shown to impede Raf/MEK kinase signal propagation (Matheny and White, 2006). Inhibition of MEK kinase has been identified as a means of maintaining pluripotency/preventing differentiation of ES cells (Ying *et al.*, 2008). Therefore, the interaction of the pluripotency factor Sall4 with this protein known to inhibit mitogenic signaling provides a valuable clue as to how Sall4 may contribute to the stabilization of the pluripotent ES cell state.

3.4.5 The Pik3cd Protein-Protein interaction network

Pik3cd (Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit delta, or p110delta) is a kinase involved in lipid metabolism and immune response regulation (Kim *et al.*, 2007). For reasons that are unknown, this isoform of PI3K is most abundantly expressed and active in both leukocytes and embryonic stem cells (Takahashi *et al.*, 2003*).* Since Pik3cd was identified in Nanog purifications with 2 peptides satisfying the high confidence interactor criteria used in this study (defined in Chapter 2: Materials and Methods, Section 2.2.3.5), reciprocal Pik3cd IP-MS/MS experiments were performed.

In these experiments, 2 peptides of Nanog were also identified in the 6XHA-Pik3cd eluates, however, they did not satisfy the high confidence criteria used in this study (defined in Chapter 2: Materials and Methods, Section 2.2.3.5).

Among the remaining 95 interactors is Fkbp3 (FK506-binding protein 3) an immunophilin which play a role in immuno-regulation and basic cellular processes involving protein folding and trafficking. This encoded protein is a cis-trans prolyl isomerase that binds the immunosuppressants FK506 and rapamycin (Ochocka *et al.*,2009) suggesting that treatment with rapamycin may alter or inhibit the characteristic metabolism conferred to ES cells because of their over-expression of Pik3cd rather than Pik3ca or Pik3cb which are over-expressed in cancer cells.

3.4.6 Integration of the protein interaction networks

The high confidence datasets of each individual network were then merged together using Cytoscape (Figure 3.20) resulting in an integrated protein-protein interaction network centered around the pluripotency factors investigated in this study. By taking this approach to visualise the data it became possible to identify key interacting proteins shared by Nanog and other protein baits. This analysis allowed for the identification of both individual proteins as well as groups of proteins (i.e., complexes) that link the Nanog interaction network to the other pluripotency factor networks.

One protein identified as connecting the Pik3cd, Sall4 and Nanog protein interaction networks is Cdc73. This protein is part of a complex (Figure 3.27) which mediates chromatin remodelling and epigenetic gene silencing (Zhu *et al.*, 2005). The identification of the Cdc73-containing chromatin remodelling/gene silencing complex connecting the Nanog, Pik3cd and Sall4 networks would suggest a mechanism of maintaining pluripotency which extends beyond simply the activation and transcription of certain genes by these pluripotency factors. That is, the

remodelling of chromatin within the nucleus which allows for the repression of genes that would cause the cells to commit to various lineages. Facilitating this process, is the Sall4-interacting protein, BRAP, which would allow for an open state of chromatin to exist in ES cells in the absence of mitogenic signalling.

As illustrated in Section 3.3.6 Integration of the Interaction Networks, the integrated interaction network reveals numerous connections between the Nanog network and the networks of the other 4 bait proteins used in this study. In addition, several connections are observed between the other 4 bait protein networks with each other. Taken together, this reveals that a multitude of contact points exist between the individual protein interaction networks and would imply that a well-integrated series of protein interaction-mediated mechanisms exist at the core pluripotency. Furthermore, the large number of proteins involved in post-translational modifications identified in this study suggests that the network of protein interactions governing pluripotency is both highly integrated and tightly regulated.

3.4.7 Comparison with the published datasets of Orkin, Van den Berg and Pardo

The protein interaction network published by Stuart Orkin's lab in 2006 (Wang *et al.*, 2006) was built around the same number of bait proteins (5 baits proteins) used in this study. The only bait protein in common with the ones used in these studies is Nanog. Comparison of the Nanog IP-LC-MS/MS presented in Section 3.3.2 with Orkin's Nanog data shows a stark contrast between Orkin's 18 Nanog interactors and the 174 determined by the Nanog IP-LC-MS/MS experiments performed in these studies. Figure 3.33 (A) shows that the overlap between the entire

Orkin published dataset and the entire dataset acquired in these studies is limited to a mere 16 proteins.

The overlap between the entire Van den Berg dataset and the dataset presented here is improved to 40. However, 15, 68 and 532 proteins remain unique to the Orkin, Van den Berg and this study's dataset, respectively. Whether these discrepancies are due to the different methods of expression employed by each study and/or the use of nuclear extracts as the starting material for the IP-LC-MS/MS experiments remains unclear.

In Figure 3.34, it is shown that although the Pardo dataset was acquired for only one bait protein (Oct4), it shows an overlap of 24 proteins. Interestingly, Figure 3.34(B) shows that the Pardo dataset also identifies CUL4B as an interactor with the protein interaction landscape for Oct4. Whether or not the use of whole cell lysates explains the identification of CUL4B both by Pardo and in this study remains to be determined. Since CUL4B can be detected both in the nucleus and the cytoplasm (Tripathi *et al.*, 2005), and Sox2 is exported to the cytosol following it acetylation (Baltus *et al.*, 2009), the IP-LC-MS/MS strategy employed here may have provided a clue as to how this protein is regulated.

3.5 Future Work

The harvesting and purification methodology developed and employed in this study provided a different view of how several key pluripotency factors are regulated post-translationally. This interaction dataset provides the field of regenerative medicine with many valuable targets for the modulation of these pluripotency factor proteins within cells.

Future work to investigate chemical compounds and small molecules that inhibit or activate the functions of the proteins listed in Tables 3.1-3.5 should be performed. Assessment and quantification of the effects of such perturbations on the identified Nanog-Sox2 interaction might provide key insights into this functional significance of this interaction. Additionally, quantification of Nanog and Sox2 protein levels in response to the inhibition of the cullin proteins identified would reveal how the protein levels of these master pluripotency may be regulated. The discovery of a molecule that could reversibly stabilize these master pluripotency factors and/or their interactions with each other studied here would provide researchers with a new-found control over the maintenance of pluripotency. Chapter 4: Exploring the Post-Translational Regulation of Pluripotency Factors Using Chemical Compounds and Small Molecules will explore this concept further.

Another major finding in this chapter is the strong links between each of the pluripotency factors and DNA damage proteins. Future work to better characterize the functional significance of these connections within mESCs should be investigated. Since DNA damage has been reported to cease under hypoxic conditions (Di Carlo *et al.*, 2004), an integrated protein-protein interaction network established under hypoxic conditions may provide insight into the dataset obtained in this chapter. Such a network may provide an even further clarified view of pluripotency protein interaction networks by answering the following: Do the interactions observed here under normoxic conditions also occur at physiologically relevant oxygen tensions? Are these interactions regulated in the same way by the same set of post-translational modifications? Do interactions between master pluripotency factors and DNA damage repair proteins still occur in the absence of DNA damage? Or, are these interactions observed solely as a consequence of a DNA

damage-inducing environment? The answers to these questions may provide valuable insight into the regulation and function of pluripotency factors as they occur *in situ*.

CHAPTER 4: Exploring the Post-Translational Modification of Nanog and Sox2 Using Chemical Compounds and Small Molecules

4.1 Introduction

Nanog levels have been shown to fluctuate in mESCs resulting in a heterogeneous population where some cells express high levels while others express very little to none (Chambers *et al.*, 2007). An initial study revealed that mESCs lacking Nanog lost pluripotency and tended to differentiate into extra-embryonic endoderm lineages upon prolonged culture (Mitsui *et al.*, 2003). Alternatively, the over-expression of Nanog completely blocks differentiation of mESCs and allows them to remain pluripotent in culture without the requirement of LIF (Chambers *et al.*, 2003; Chambers *et al.*, 2004). As a result of these combined properties, Nanog has become regarded as being 'the gate-keeper' of pluripotency.

In the years following, it has been well-documented that Nanog protein levels are modulated both by phosphorylation which causes self-dimerization and stabilizes it and also by ubiquitination which results in its degradation by the 26S proteasome (Mullin *et al.*, 2008; Wang *et al.*, 2008; Moretto-Zita *et al.*, 2010; Ramakrishna *et al.*, 2011; Chae *et al.*, 2012). However, neither the kinase phosphorylating Nanog nor the E3 ligase(s) ubiquitinating it and targeting it for proteasomal degradation have been identified to date.

The identification of chemical compounds or small molecules that could elevate the levels of phosphorylated Nanog thereby, stabilizing its transcriptional activity would provide a non-genetic means of blocking differentiation. One approach to elevating Nanog protein levels in cells would be to identify the E3 ligase

responsible for its degradation and inhibit its function. Recently, a small molecule named MLN4924 has been reported as an inhibitor of the enzyme NAE1, efficiently blocking the Neddylation of all cullin proteins and blocking the degradation of CRL substrate proteins (Soucy *et al.*, 2009). If the master regulators of pluripotency were found to be substrates of CRLs, MLN4924 may provide a reversible and dose-dependent method of stabilizing pluripotency. Understanding how proteins that confer pluripotency to ES cells are both stabilized and degraded will reveal the biochemical mechanisms by which the stem cell state is established, maintained, and eventually lost. The ability to control and modulate the stability of pluripotency itself through the use of small molecules could provide researchers with valuable tools for future fundamental research, as well as, potentially advance the field of regenerative medicine.

Conversely, the identification of the kinase responsible for phosphorylating Nanog and inhibition of its function would result in opening the door toward differentiation. Such a compound may allow for all cells in a given culture to be homogeneously primed for differentiation. These cells could then be differentiated into the lineage of choice by subsequent treatments. This could provide a powerful reagent for cell-based therapeutics to treat degenerative diseases.

In addition to ES cells, Nanog has also been detected in several tumour cell types including germ cell tumours, testicular, lung, oral, gastric and prostate cancers (Hoei-Hansen *et al.*, 2005; Hart *et. al.*, 2005; Jeter *et al.*, 2009; Chiou *et al.*, 2008; Chiou *et al.*, 2010). It has been found to play a key role in tumour development by conferring pluripotent stem cell characteristics to tumour cells (Jeter *et al.*, 2011; Chiou *et al.*, 2010). In the field of oncology, Nanog is currently being used as a

diagnostic biomarker. Its presence is indicative of the most invasive forms of cancer and a poor prognosis (Meng *et al.*, 2010; Jeter *et al.*, 2011; Tsai *et al.*, 2011). Therefore, the discovery of a chemical compound or small molecule that could prevent the phosphorylation of Nanog and thereby destabilize it would induce differentiation of highly metastatic cancer cells. This compound would result in the development of new reagents to combat the most aggressive types of cancer.

4.1.1 Proteins involved in post-translational modifications identified in this study

In Chapter 3, protein-protein interaction networks were obtained for five different pluripotency factors. Among the identified interactors, were several proteins responsible for performing various types of post-translational modifications and they are summarized in Table 4.1.

With regards to the degradation of proteins, the cullin proteins that have been identified in this study to potentially interact with pluripotency factors are of particular interest. CUL3 was identified in Sox2 purifications, CUL4B was identified in both Nanog and Sox2 purifications and CUL7 was identified to co-purify with both Nanog and Esrrβ.

	Bait Protein:				
Protein Class:	**Nanog**	**Sox2**	**Esrrβ**	**Sall4**	**Pik3cd**
Kinases	Pik3cd	ATM	AurkA	Riok1	Pik3r3
	FRAP1	ERK1	Skiv2l		Pik3r2
	Sphk2	BUB1	Riok2		Ak2
	Mastl	Melk			
	Oxsr1	Cit			
	Prkar1a	Chka			
	Cdk7	Sgk223			
Cullins	CUL4B	CUL4B	CUL7		
	CUL7	CUL3			
E3 ligase	Trim33		Ubr5	Ubr5	Stub1
	Rnf2			Rnf2	
	Huwe1	Huwe1		Mib2	
	Dtx2			MycBP2	
	Cdc23			BRAP	
E2 conjugating		Ube2o			Ube2i
		Ube2s			
Ub peptidase		Usp28			
		Usp34			
SUMO E3 ligase			RanBP2		

Table 4.1: Proteins involved in post-translational modifications identified in this study. Summary of all proteins involved in post-translational modifications identified in all 5 protein-protein interaction networks.

4.1.2 Chemical compounds and small molecules used to study the regulation of Nanog and other pluripotency factors

As highlighted in Table 4.1, it is possible that multiple pluripotency-conferring proteins are targeted for degradation by different CRLs. Further exploration and identification of which CRL is degrading which protein is necessary in order to reveal how these proteins might be stabilized and their abundances controlled without the need for genetic manipulation.

This Chapter will investigate the perturbation of the protein modifying enzymes identified in Chapter 3 using a variety of molecules purposefully chosen based on the identification of their previously published target proteins by IP-LC-MS/MS in this study. A summary of the molecules selected for use in this Chapter are listed and their properties summarized in Table 4.2.

Name of Molecule	Empirical Formula	Molecular Weight	Known targets	Target(s) identified by IP-LC-MS/MS
H89	$C_{20}H_{20}BrN_3O_2S \cdot 2HCl \cdot H_2O$	519.28	PKA PKB/Akt1 PKC Prkar1a Prkar1b Prkar2a Prkar2b	AKAP8 AKAP8l Prkaca (PKA catalytic subunit) Prkar1a (PKA regulatory subunit)
Forskolin	$C_{22}H_{34}O_7$	410.50	Adenylyl cyclase Activator, increases intracellular cAMP levels	Prkaca (PKA catalytic subunit) Prkar1a (PKA regulatory
Wortmannin	$C_{23}H_{24}O_8$	428.43	All Pi3 Kinases and the Pi3K/Akt cascade	Pik3cd Pik3r3 Pik3r2 Irs2
MLN4924	$C_{21}H_{25}N_5O_4S$	443.52	Nedd8-activating enzyme inhibitor. Inhibits all CRL activity. Arrests cells in S-phase.	Nedd8 CUL3 CUL4B CUL7
Ht-31	Stearyl-DLIEEAASRIVDAVIEQVKAAGAY amide	1320	The AKAP-binding domain of Prkar2a. Inhibits the interaction between PKA regulatory subunit RII with all AKAPs.	AKAP8 AKAP8l
PKI (14-22) amide	$C_{53}H_{98}N_{20}O_{12} \cdot C_2HF_3O_2$	1321.50	PKA	Prkaca (PKA catalytic subunit)
Br-cAMP	$C_{10}H_{11}BrN_5O_6P$	408.10	Increasing cellular cAMP levels	Prkaca (PKA catalytic subunit)
3*i* ** [0.8 μM] PD184352	$C_{17}H_{14}ClF_2IN_2O_2$	478.66	PD184352: MEK-1	MEK-1
[2 μM] SU540	$C_{17}H_{16}N_2O_3$	296.32	SU5402: FGF-R1 resulting in inhibition of phospho-ERK1 & 2	Erk1 & 2
[3 μM] CHIR99021	$C_{22}H_{18}Cl_2N_8$	465.34	CHIR99021: GSK-3β	GSK-3β
Aurk*i* III	$C_{23}H_{20}N_4O_3$	400.4	AurkA AurkB	AurkA
Nocodazole	$C_{14}H_{11}N_3O_3S$	301.32	Inhibits microtubule dynamics. Arrests cell cycle at the G2/M phase.	Used here to synchronize cells in Mitosis

Table 4.2: Chemical compounds and small molecules used to study the regulation of Nanog and Sox2. Summarized information regarding the structural, chemical and functional properties of compounds used in this study. Data presented here has been extracted from the manufacturers websites with the exception of Ht-31 for which some information can be found under US patent number 6610657 (Promega Corp.).

*****3*i* inhibitor cocktail as described by Ying *et al.*, 2008. Surpassed by 2*i* (MEK-1/2 and GSK-3β) inhibitor cocktail by Silva *et al.*, 2008.

4.2 Aims of this chapter

The aim of the work in this chapter was to further explore the post-translational protein modifying enzymes identified in the previous chapter to interact with the five pluripotency factors characterized in this study. The combined approach of chemical compound/small molecule inhibition and native gel analysis of Nanog protein complexes was employed to investigate the signaling pathways controlling the stability of Nanog-containing protein complexes. Furthermore, this approach will provide a means to assess the stability of the Nanog interaction with Sox2 and determine its involvement in maintaining the pluripotent state. Co-localization studies will be performed in an attempt to visualize whether or not the interactions observed by IP-LC-MS/MS in this study (Chapter 3) naturally occur within the intact cells. The development of several multiplexed fluorescence based quantitative confocal microscopy assays, will allow for the determination of Nanog and Sox2 protein levels in response to various perturbations including the inhibition of CRLs. The use of the NAE1 inhibitor MLN4924 will be used, as well as, shRNA knockdown of CUL3, CUL4B and CUL7 in order to identify the CRL that may responsible for targeting Nanog and Sox2 for degradation. Confocal microscopy will be used to examine whether or not CUL4B co-localizes with Nanog and Sox2 and whether or not this interaction occurs within the cytoplasm or the nucleus. This will address the issue raised in Chapter 3 about whether or not IP-LC-MS/MS studies performed on whole cell lysates accounts for the identification of the CUL4B protein in both the dataset reported by Pardo *et al.* in 2010 and in this study. Furthermore, the effects of inhibiting CUL3, CUL4B and CUL7 on the cell cycle of mESCs will

be studied in order to determine at which point in the cell cycle Nanog is being targeted for degradation.

4.3 Results:

4.3.1 Effects of various chemical compounds and small molecules on the Nanog protein

In this experiment, the 3XFLAG-Nanog stable cell line was treated with the inhibitors listed in Table 4.2 in order to determine if their respective targets were responsible for the phosphorylation of Nanog protein in mESCs (Yates and Chambers, 2005).

In order to confirm that this increased molecular weight band detected by the anti-FLAG M2 antibody was indeed a phosphorylated form of the Nanog protein, a λ-phosphatase treatment was performed on immunoprecipitated 3XFLAG-Nanog protein complexes. In Figure 4.1 (A), it is shown that treatment with λ-phosphatase results in the disappearance of the higher molecular weight band detected by the anti-FLAG M2 antibody.

In Figure 4.1 (B), an increase in phospho-Nanog levels is observed following a 12 hour treatment of mESCs with PKI (14-22), Ht-31, MLN4924, Aurk*i*III and nocodazole. In comparison, 3*i* cocktail caused only a marginal increase in phospho-Nanog levels. Conversely, the band corresponding to phospho-Nanog completely disappears following treatment with H89, Forskolin and Wortmannin. Interestingly, treatment with Nocodazole, known to synchronize cells at the G2/M stage of the cell cycle causes the accumulation of phospho-Histone H3 as well as an increase in phospho-Nanog levels.

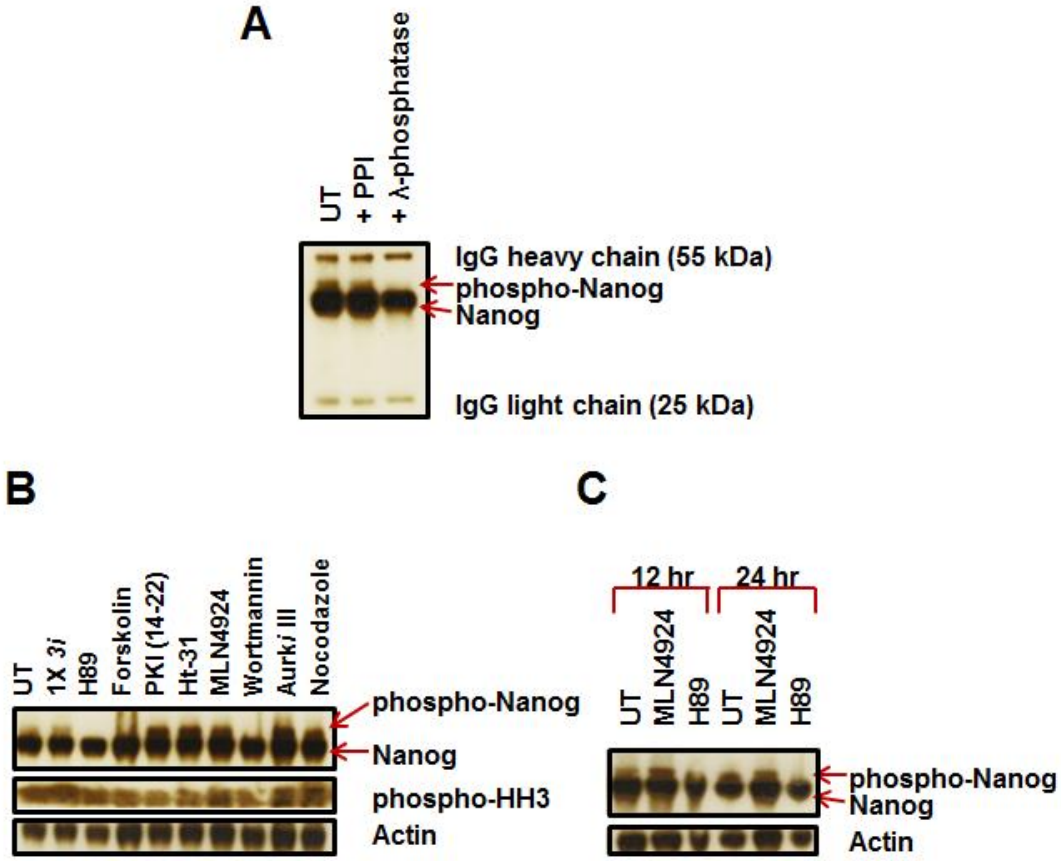


Figure 4.1: Effects of various chemical inhibitors on the Nanog protein in mouse embryonic stem cells. A) Western blot on lysates prepared from 3XFLAG-Nanog stable cell line treated with various inhibitors for 12 hours. Arrows point to the Nanog protein and a modified Nanog. **B)** anti-FLAG western blot on FLAG M2 agarose- purified 3XFLAG-Nanog protein. Nanog protein bound to beads were either untreated (UT) or subjected to phosphatase inhibitor (PPI) and λ-phosphatase treatment. Arrows point to Nanog and phospho-Nanog proteins. **C)** Effects of the compounds H89 and MLN4924 on Nanog phosphorylation status after 12 and 24 hours of treatment. Arrows indicate Nanog and phospho-Nanog bands. anti-phospho-HH3 was used as a mitotic marker and anti-actin as a loading control.

In Figure 4.1 (C), it is shown that the opposing effects of MLN4924 and H89 treatment on the abundance of phospho-Nanog levels are consistent after 12 and 24 hours of treatment in this study.

4.3.2 Effects of various perturbations on Nanog-containing protein complexes

In order to determine the how the disappearance of phospho-Nanog affects the formation of Nanog complexes, 3XFLAG-Nanog complexes were immunopurified from mESCs cultured in normoxia, at physiologically relevant

oxygen tensions treated with H89 and separated by native gel electrophoresis.

Figure 4.2 (A) shows that the molecular weight of the most prominent 3XFLAG-Nanog-containing complexes decreases from well above 1.2 MDa to between 1 MDa and 0.75 MDa in cells treated with H89. Furthermore, the complexes appear as a relatively uniform smear across the molecular weight range indicating a lack of resolution by the native gel of assembled 3XFLAG-Nanog complexes. Figure 4.2 (B – D) shows the distribution of 3XFLAG-Nanog complexes after two-dimensional Native-to-SDS PAGE. In ES cells grown in standard culture conditions (B), a broad range in the distribution of complexes exists with the most abundant complexes having a molecular weight of between 1 MDa and 0.5 MDa. At physiologically relevant oxygen tensions (C), a significant decrease in modified 3XFLAG-Nanog was observed. Strikingly, in Figure 4.2 (D), following H89 treatment those complexes within the 1 MDa to 0.5 MDa size range were no longer the most abundant complexes in mES cells. Instead, those complexes between 0.242 MDa and 0.040 MDa became both the most abundant.

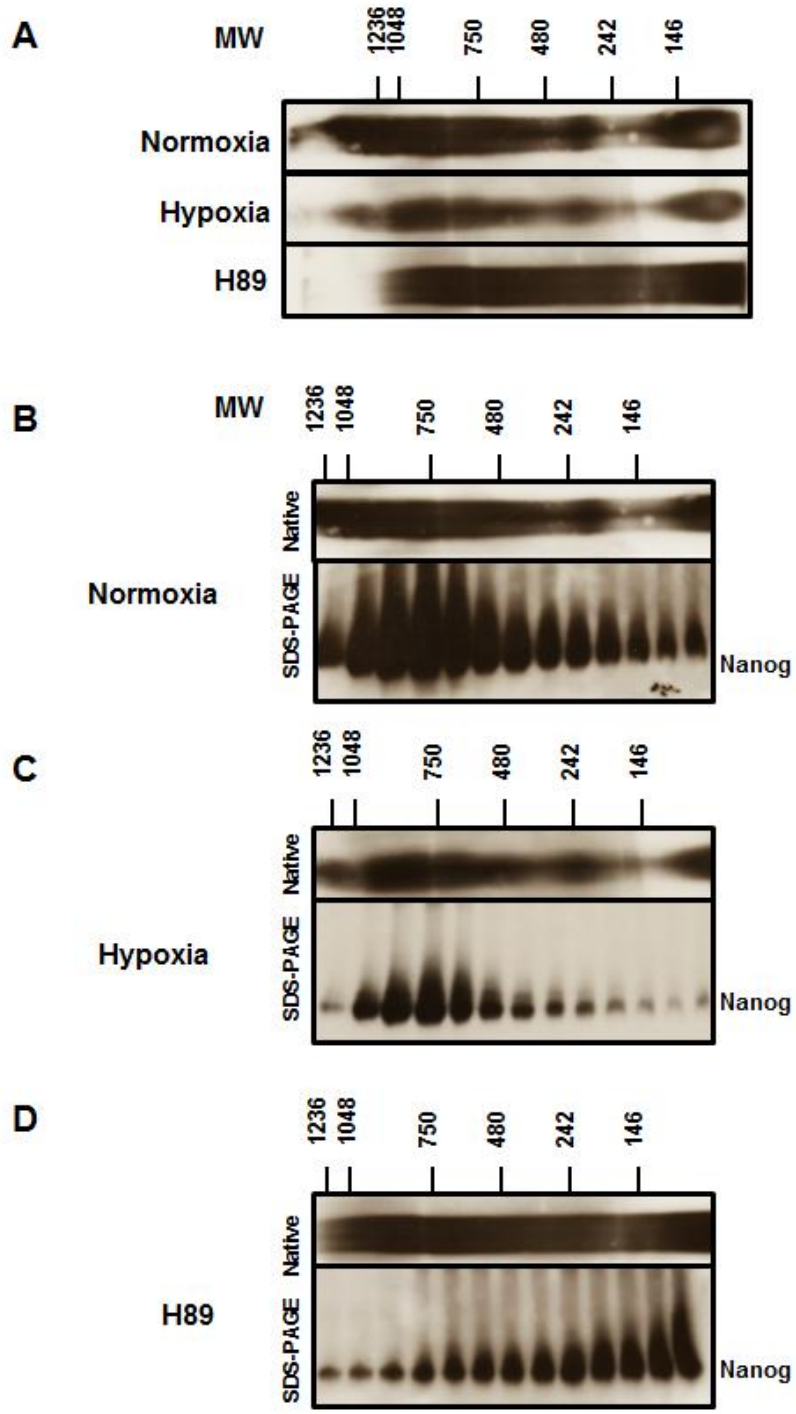


Figure 4.2: Native and Native-to-SDS-PAGE western blot analysis of Nanog containing protein complexes. anti-FLAG western blot performed on M2 agarose purified 3XFLAG-Nanog protein complexes isolated from mES cells subjected to various perturbations and separated on **A)** NativePAGE gradient native gel. NativePAGE-to-SDS-PAGE analysis was performed on 3XFLAG-Nanog-containing complexes isolated from mES cells cultured in: **B)** Normoxia, **C)** 24 hours hypoxia (5% O_2), and **D)** 24 hours treatment with 10μM H89.

In order to examine how the elevation of phospho-Nanog affects the formation of Nanog complexes, native gel analysis was performed on

immunopreciptates from cells under various perturbations, including treatment with MLN4924 and Ht-31.

Figure 4.3 shows a comparison of the molecular weight distribution of 3XFLAG-Nanog complexes in response to various forms of phospho-Nanog stabilization. The most abundant Nanog complexes are within the range of >2.0 MDa to 0.75 MDa in size. The effects of the various chemical treatments are apparent at the lower molecular weight range of the gel corresponding to the size of the dimeric form of Nanog. Here we see that culture in hypoxic conditions as well as the small molecules Ht-31 and Forskolin increase the levels of dimeric Nanog whereas MLN4924 and Wortmannin (a PI3K inhibitor) seemingly decrease the levels of dimeric Nanog.

It was then determined what effect phospho-Nanog and possibly Nanog dimerization has on Nanog protein-protein interactions in ES cells. Since Sox2 was identified as an interactor with Nanog by the IP-LC-MS/MS studies performed in Chapter 3, the interaction between these two proteins in the presence and absence of phospho-Nanog was examined. In Figure 4.4 (A), shown is a representative native gel analysis of the molecular weight distribution of Nanog protein complexes in response to various inhibitor treatments. (B) to (G) show the distribution of Nanog and Sox2-containing protein complexes in mESCs after the same treatments subjected to Native-to-SDS PAGE analysis.

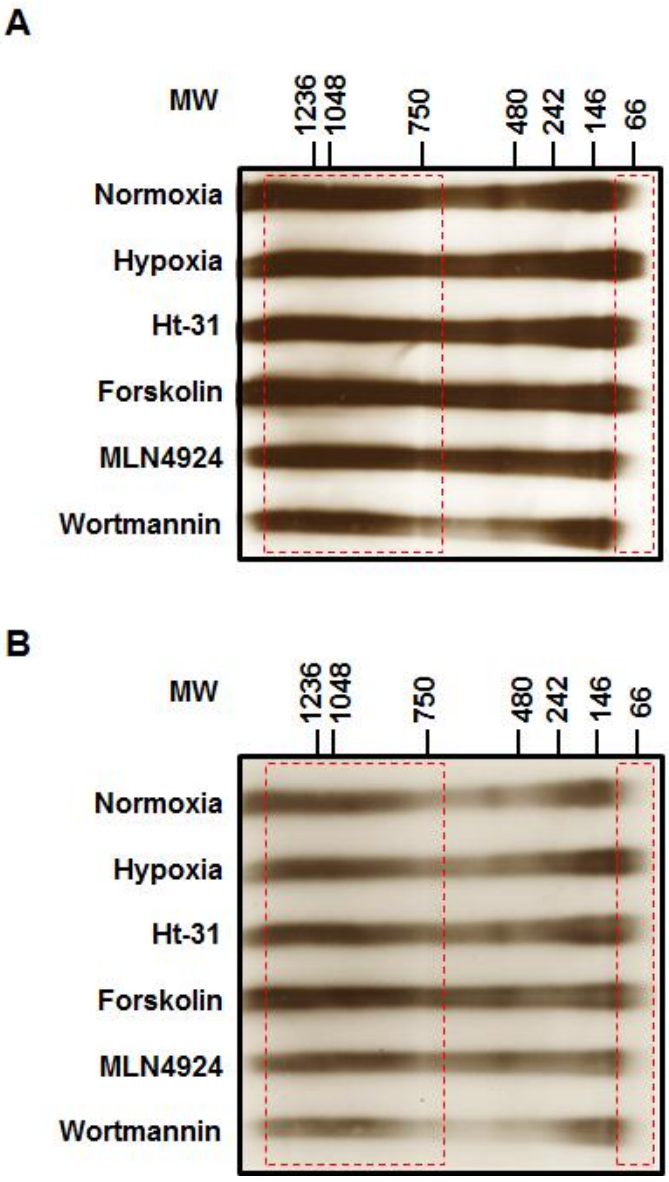


Figure 4.3: Size distribution of Nanog-containing protein complexes in response to various perturbations. anti-FLAG western blot on BN-PAGE separated immunopurified 3XFLAG-Nanog-containing complexes in response to various perturbations. **A)** 30 second exposure time and, **B)** 10 second exposure time of the same gel. Red boxes outlining both the most abundant species of Nanog complexes (75-1500 kDa) and possibly the dimeric form of Nanog (approx. 68 kDa).

In Figure 4.4 (A), an anti-FLAG (Left-hand panels) and an anti-Sox2 (Right-hand panels) western blot on BN-to-SDS-PAGE separated 3XFLAG-Nanog complexes are shown. Figure 4.4 (B), shows that in standard culture conditions, Nanog and Sox2 interact only in the highest molecular weight complexes near the resolution limit of the gel, approx. 20 MDa. (C) Shows that the protein complex distribution profile in physiologically relevant conditions is exactly the same as normoxia but has a decreased mean molecular weight between 1.2 to 1.0 MDa. (D) Shows that treatment with Ht-31 results in all Nanog complexes containing the Sox2 protein. These complexes have a mean molecular weight distribution similar to hypoxia, between 1.2 to 1.0 MDa. (E) Shows that treatment with Forskolin results in Nanog and Sox2-containing protein complexes forming at > 1.2 MDa to near 20 MDa (the resolution of the gel). In addition, all complexes are heavily modified in this condition. (F) Shows that treatment with MLN4924 results in all complexes containing both Nanog and Sox2 across the molecular weight range observed. In addition, the protein complexes in the mean molecular weight range (>1.2 MDa to 0.8 MDa) contain equal amounts of both Nanog and Sox2 protein. (G) Suggests that treatment with Wortmannin potentially induces the degradation of Nanog.

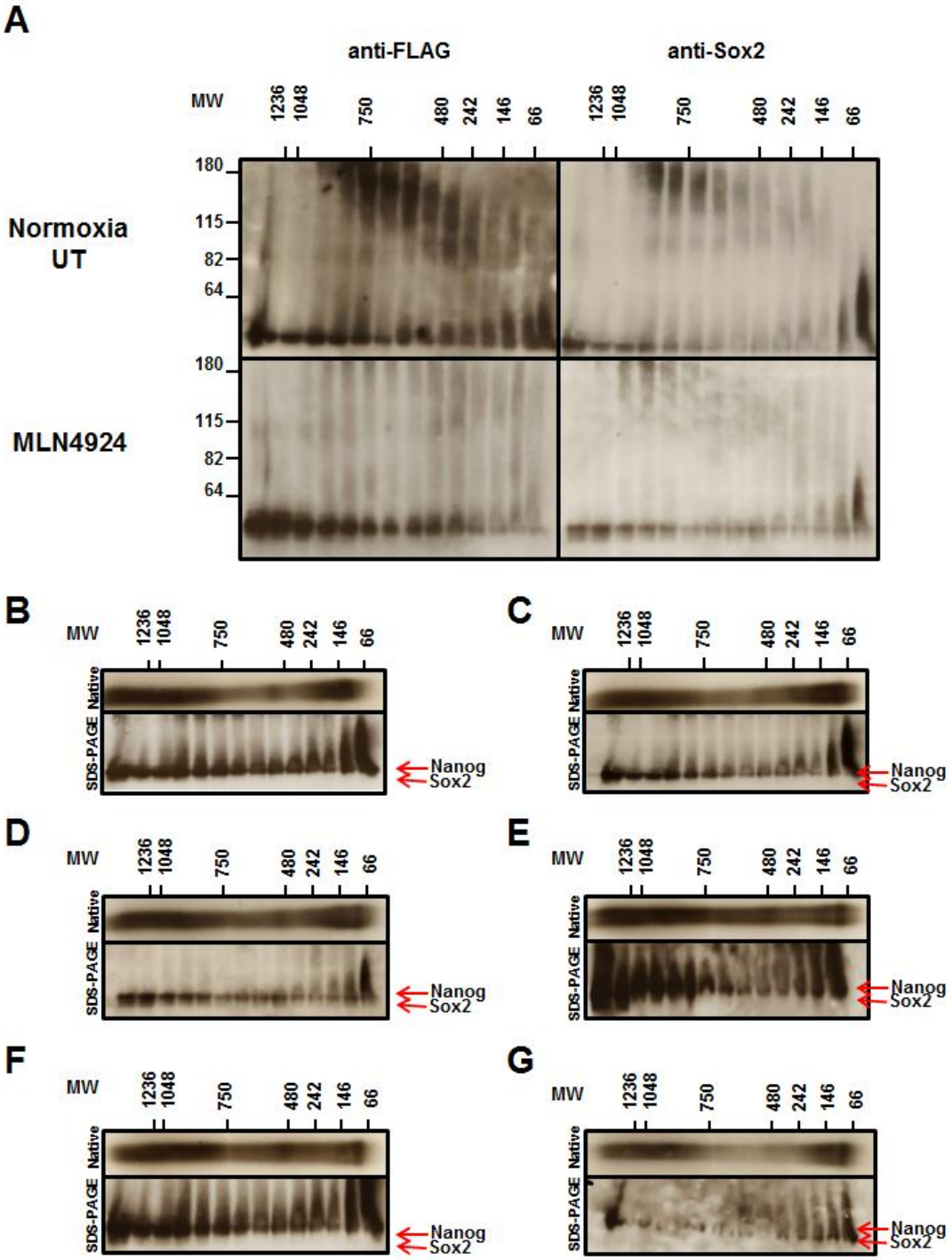


Figure 4.4: The effects of chemical compounds and small molecules on the interaction between Nanog and Sox2. A) anti-FLAG and anti-Sox2 western blotting of BN to SDS-PAGE separated immunopurified 3XFLAG-Nanog-containing protein complexes. BN to SDS-PAGE western blotting for Nanog (anti-FLAG) and anti-Sox2 of protein complexes isolated from 3XFLAG-Nanog stable cell line cultured for 12 hours in: **B)** Normoxia, **C)** Hypoxia, **D)** 50 µM Ht-31, **E)** 50 µM Forskolin, **F)** 10 µM MLN4924 and, **G)** 50 µM Wortmannin.

4.3.3 Quantitation of Nanog, Sox2 and CUL4A/B proteins in response to inhibition of the PKA signaling pathway

In this experiment the levels of Nanog, Sox2 and CUL4A/B proteins in the entire cell was measured in response to various perturbations. 3XFLAG-Nanog mESCs subjected to various treatments were triple immunostained for all three proteins. Figure 4.5 shows overlay images of 3XFLAG-Nanog, endogenous Sox2 and endogenous CUL4A/B proteins in mESCs. While the 3XFLAG-Nanog protein is predominantly nuclear, Sox2 and CUL4A/B proteins are mostly cytoplasmic. The Sox2 protein is distributed through both the cytoplasm and the nucleus, the CUL4A/B signal is detected almost exclusively in the cytoplasm.

Relative fluorescence measurements were extracted from raw confocal data using Macro1 (see APPENDIX F). Briefly, this Macro quantifies the relative fluorescence units (RFUs) in the red channel, as thresholded by the user and maps the reading from the green channel onto the areas of fluorescence measured in the red channel. The Macro performs this function on every slice in the entire z-stack of images acquired. The user assigns which fluorescence channel (green, red or far red) is to be measured as the red and green channels defined by the macro script.

In Figure 4.6 (A), the RFU measurements were extracted with Nanog being assigned as the red channel and CUL4A/B assigned as the green. In Figure 4.6 (B), the RFU measurements were extracted with Sox2 being assigned as the red channel and CUL4A/B assigned as the green. Data was treated as a multiplexed fluorescence-based assay where the mean intensities averaged from 3 z-stacks were converted into percentage relative fluorescence. Therefore, the results are represented in terms of a percentage increase or decrease relative to untreated control samples.

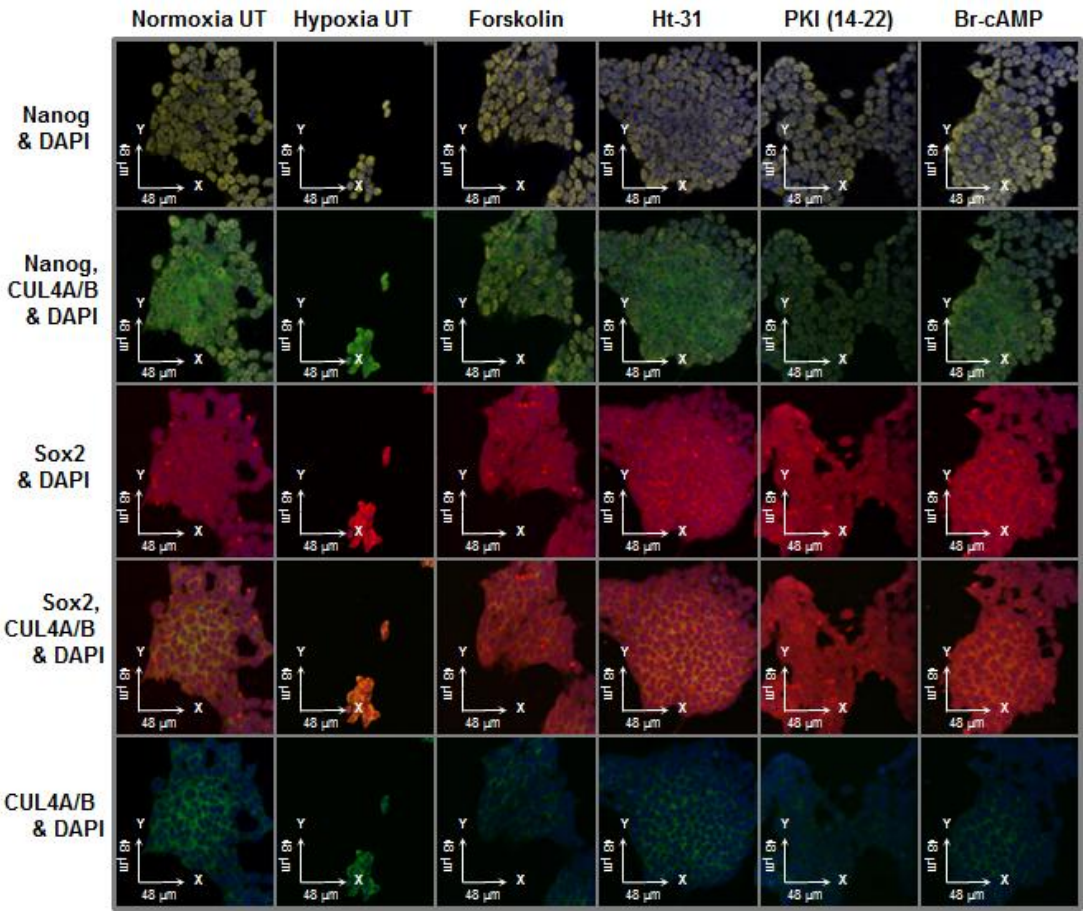


Figure 4.5: Immunostaining for Nanog, Sox2 and CUL4A/B in mESCs subjected to various perturbations. Immunofluorescent overlays of the 3XFLAG-Nanog stable cell line subjected to various treatments for 12 hours and immunostained for Nanog (yellow), Sox2 (red) and CUL4A/B (green). Scale bar is 48 µm.

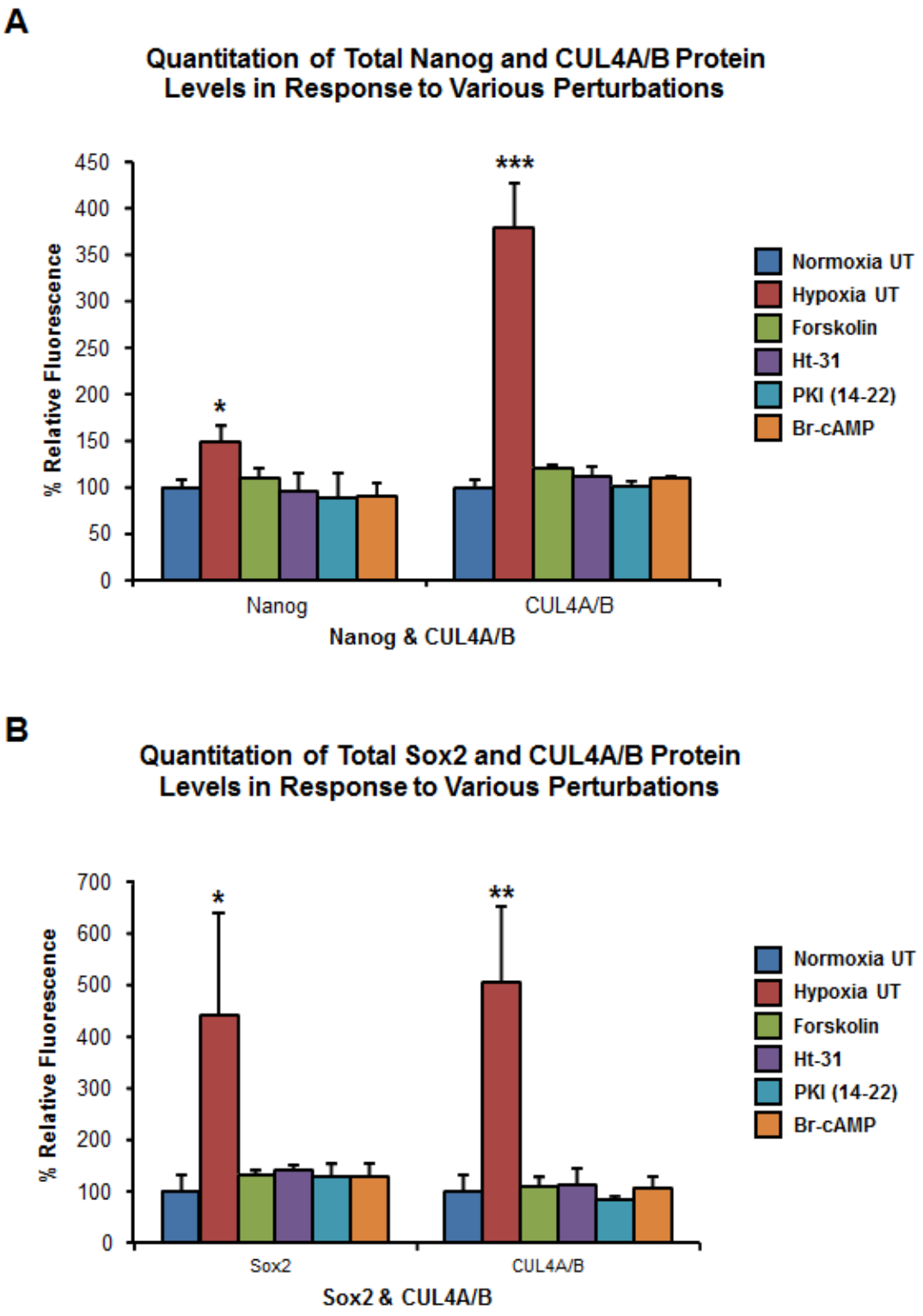


Figure 4.6: Effects of chemical inhibition on total cellular Nanog, Sox2 and CUL4A/B protein levels. A) Total cellular Nanog relative to CUL4A/B protein levels. **B)** Total cellular Sox2 relative to CUL4A/B protein levels. For all graphs, the relative fluorescence averaged from 3 separate mESC colonies was calculated using Normoxia UT Control fluorescence as the denominator. 1 St. Dev. above the mean % relative fluorescence is shown. The significance compared to Normoxia UT control was determined using a one-way ANOVA test and is indicated with asterisks above error bars. Where * = p-value < 0.05 and ** = p-value < 0.01.

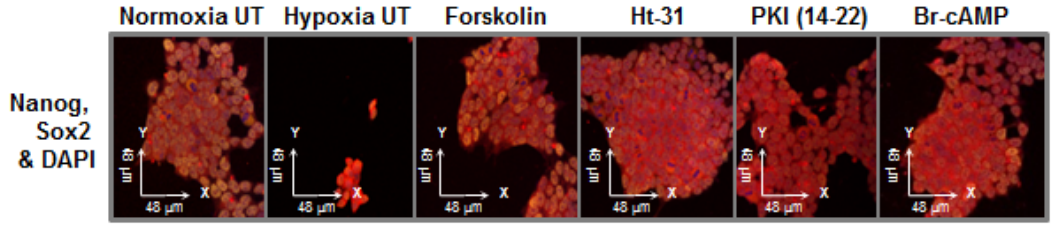
Figure 4.6 (A) Shows the quantitation of total Nanog protein levels relative to the total CUL4A/B protein levels. Of all 5 treatments tested in this experiment, only hypoxia resulted in an increase in either total Nanog or total CUL4A/B protein levels by 49% and 280%, respectively.

Figure 4.6 (B) Shows the quantitation of total Sox2 protein levels relative to the total CUL4A/B protein levels. Once again, only hypoxia revealed an effect on the protein levels of either Sox2 or CUL4A/B. That is, an increase of 341% in Sox2 protein levels and an increase of 406% in CUL4A/B protein levels.

In order to quantify nuclear Nanog protein levels relative to Sox2 protein levels, data was extracted from the same set of z-stack images shown in Figure 4.5 using Macro3 (see APPENDIX G). Briefly, once the DAPI signal has been thresholded by the user, this Macro traces outlines around the DAPI signal (nuclear area) and measures the fluorescence intensity (RFU) in the green channel mapped onto the DAPI signal. The extraction method was performed in two steps. The first extraction with the Nanog channel (far red) set as the green channel and the second one where the Sox2 channel (red) was set as the green channel.

Figure 4.7 (A) Shows overlay images of Nanog and Sox2 proteins following various treatments. (B) Shows the quantitation of nuclear Nanog and Sox2 protein levels, as well as, nucleus area measurements in response to the various treatments. No significant changes in nuclear Nanog protein levels were observed in any of the conditions tested. An increase in nuclear Sox2 levels was detected in response to hypoxia, Ht-31 and PKI (14-22) amide treated cells by 87%, 38% and 38%, respectively. Nucleus area was decreased by 30% in hypoxia and by 14% by Ht-31 treatment.

A



B

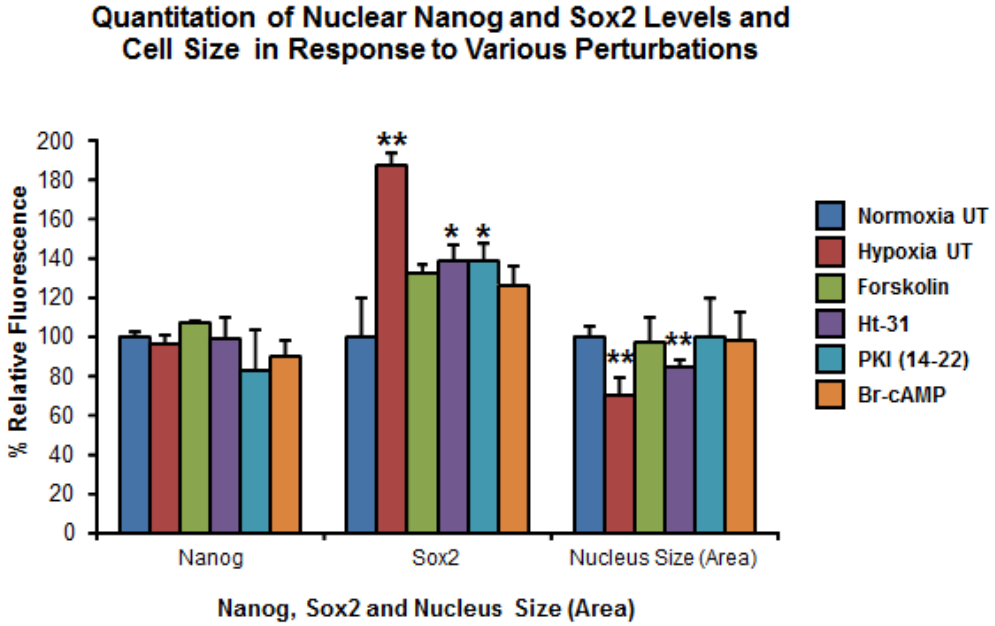


Figure 4.7: Effects of chemical inhibition on nuclear Nanog and Sox2 protein levels in function of nucleus size. A) Immunofluorescent image overlay of Nanog (yellow), Sox2 (red) and DAPI (blue). Scale bar is 48 µm. **B)** Nuclear Nanog relative to nuclear Sox2 protein levels and nucleus size. For all graphs, the relative fluorescence was calculated using Normoxia UT Control fluorescence as the denominator. The significance compared to Normoxia UT control was determined using a one-way ANOVA test and is indicated with asterisks above error bars. Where: * = p-value < 0.05 and ** = p-value < 0.01.

4.3.4 Co-localization of Nanog, Sox2 and CUL4A/B proteins

In this experiment the co-localization co-efficients of Nanog, Sox2 and CUL4A/B protein were determined using Volocity 6.0.

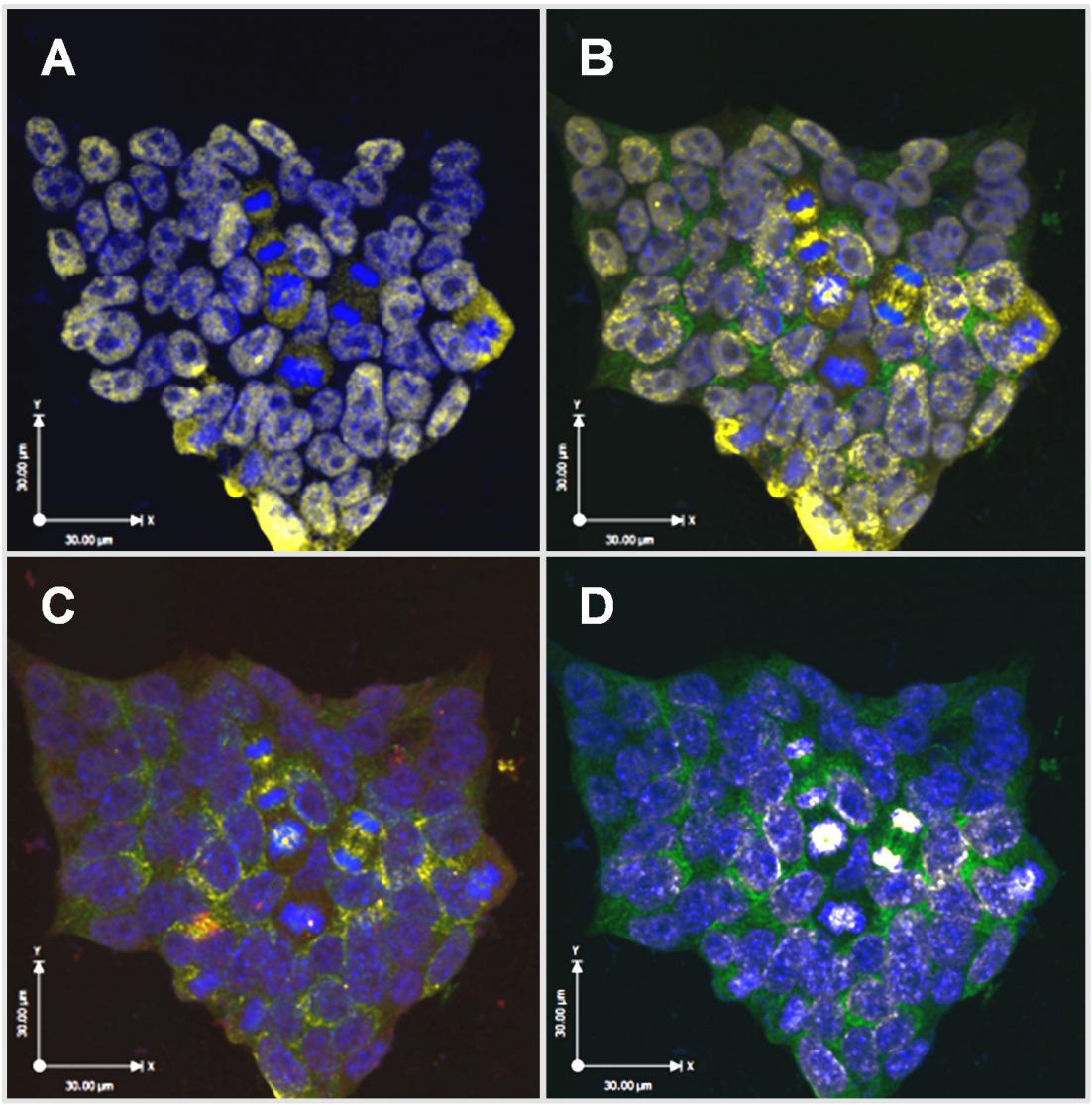


Figure 4.8: Cellular localization and co-localization of Nanog, Sox2 and CUL4A/B proteins in mouse embryonic stem cells. (A) Nanog and Sox2 co-localization (Pearson's = 0.643), **(B)** Nanog and CUL4A/B co-localization (Pearson's = 0.439), **(C)** Sox2 and CUL4A/B co-localization (Pearson's = 0.738) and **(D)** CUL4A/B and DAPI co-localization (Pearson's = 0.472). Scale bar is 30 µm. Co-localizing pixels are shown in yellow, CUL4A/B in green, Sox2 in red and DAPI blue.

In Figure 4.8 (A) Shown is Nanog and Sox2 protein co-localization in the nucleus. (B) Shows Nanog and CUL4A/B protein co-localization in the nucleus. Co-localization was increased at mitotic spindles as chromosome segregation occurred during anaphase and telophase. (C) Shows co-localization between Sox2 and

CUL4A/B occurring in the cytoplasm. A dramatic change was observed during anaphase and telophase where they co-localized in the nucleus at mitotic spindles. (D) Shows a poor co-localization between CUL4A/B and the DAPI signal until mitosis where they co-localized strongly, on condensed chromosomes.

4.3.5 Effects of the chemical compound MLN4924 on the protein levels of various pluripotency factors

In this experiment, the 3XFLAG-Nanog mES stable cell line was treated with the chemical compound MLN4924 in order to determine if inhibiting CRL activity would increase the levels of Nanog, Sox2 and Esrrβ proteins. Figure 4.9 (A) shown is an anti-FLAG western blot demonstrating an increase in the levels of phospho-Nanog following 12 and 24 hours of CRL inhibition compared to untreated cells and H89-treated cells. (B) Shown is an anti- Esrrβ western blot performed on mESCs treated for 12 and 24 hours with MLN4924. Inhibition of CRL function did not have any effect on Esrrβ protein levels. (C) Shown is a western blot to detect Sox2 and CUL4A/B protein levels in response to CRL inhibition after 12 and 24 hours. MLN4924 treatment resulted in increases in both Sox2 and CUL4A/B protein levels over time.

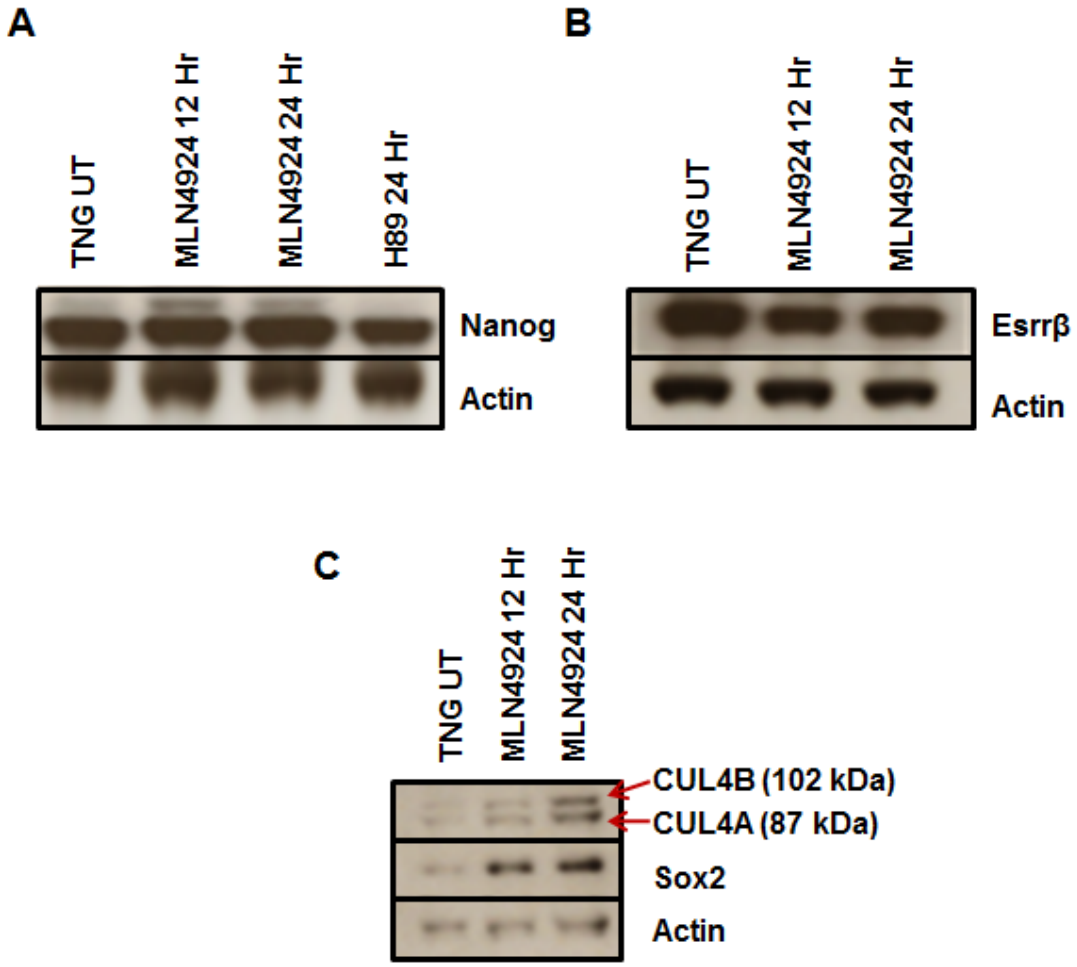


Figure 4.9: Effects of inhibiting Cullin Ring Ligase (CRL) activity with the compound MLN4924 on Nanog, Sox2 and Esrrβ proteins. A) anti-FLAG western blot showing Nanog and phospho-Nanog proteins in the 3XFLAG-Nanog stable cell line treated with MLN4924 versus H89 treatment. **B)** anti-Esrrβ western blot showing levels of Esrrβ protein in response to cullin inhibition. **(C)** Western blot showing Sox2 and CUL4A/B proteins following treatment with MLN4924. Anti-Actin was used as a loading control in these experiments.

4.3.6 Effects of individual cullin-specific shRNA on the interaction of Nanog with Sox2 and Esrrβ

The increase in phospho-Nanog caused by the pan-Cullin inhibitor MLN4924 was further dissected using a more targeted cullin inhibition approach. To this end, stable cell lines were created expressing shRNA targeted to CUL3, CUL4B, CUL7, as well as, a scrambled control sequence. In order to examine how this increase in phospho-Nanog affects Nanog protein interactions, 3XFLAG-Nanog protein was immunopurified from the shRNA-expressing stable cell lines and examined for the presence of Nanog, Sox2 and Esrrβ.

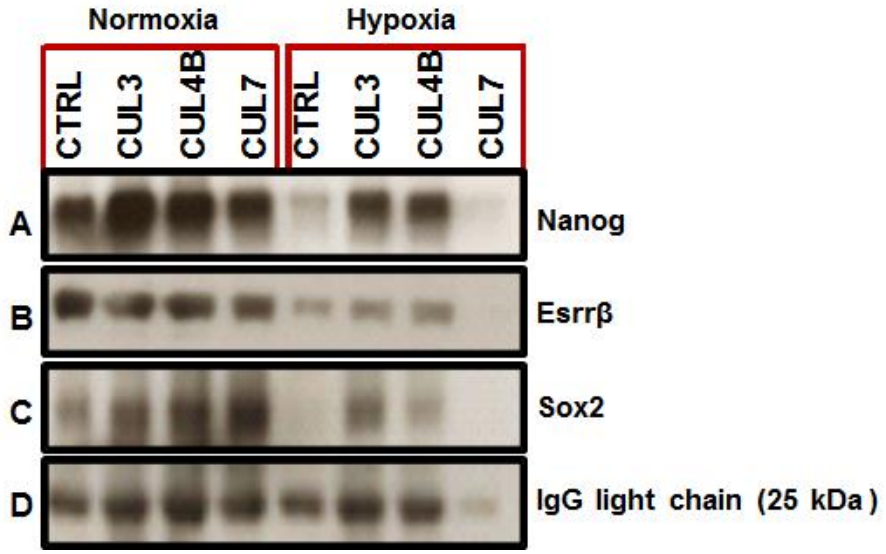


Figure 4.10: The effect of specific knockdown of individual cullins on Nanog interaction with Sox2 and Esrrβ. (A) anti-FLAG IP western blot showing Nanog protein levels purified from ES cells in response to the inhibition of CUL3, CUL4B and CUL7 proteins **(B)** anti-Esrrβ blot showing the levels of Esrrβ protein co-purifying with Nanog protein in response to cullin inhibition**. (C)** anti-Sox2 western blot showing the levels of Sox2 protein co-purification with Nanog protein in response to culling inhibition, and **(D)** levels of IgG light chain used as a loading control.

Figure 4.10 (A) Shows that knockdown of CUL3, and to a slightly lesser extent, knockdown of CUL4B resulted in an increase in the level of Nanog being immunoprecipitated. (B) Shows the amount of Esrrβ that co-purified with Nanog remained unaffected by the knockdown of CUL3, CUL4B or CUL7. (C) Shows the amount of Sox2 protein co-purified with Nanog increased in response to all cullin

knockdowns and was highest in response to CUL4B and CUL7 shRNA in normoxia. Unexpectedly, after 16 hours in hypoxic conditions, Nanog was barely detectable in control cells and both Sox2 and Esrrβ completely disappeared in CUL7 shRNA.

4.3.7 Quantitation of Nanog and phospho-Histone H3 protein levels in response to cullin inhibition

In order to confirm whether or not the increase in the amount of Nanog protein purified from CUL3 and CUL4B knockdown cell lines was due to an increase in Nanog protein levels in these cells, a multiplexed quantitative confocal microscopy assay was performed. In this experiment Nanog protein levels were measured in response to various forms of cullin inhibition. Since CUL3 inhibition has been reported to delay entry into mitosis, as well as, increase in the levels of intracellular reactive oxygen species (ROS), both ROS and the phospho-Histone H3 (pHH3) levels were quantified as well. All image analysis was performed using Volocity 6.0.

Figure 4.11 shows the fluorescent overlay images of Nanog (shown in red), pHH3 (shown in green) and ROS (shown in yellow). In these images it is shown that treatment with MLN4924 causes an apparent increase in nucleus size accompanied by a lack of pHH3 staining. Another interesting observation made in this experiment was that the ROS staining was primarily restricted to the cytoplasm in all mESC cell lines examined.

In order to quantify the levels of Nanog, pHH3 and ROS from these images, the mean fluorescence intensities averaged from 3 independent colonies was extracted using Volocity 6.0 and the relative percentage fluorescence was calculated

using CTRL shRNA as the denominator. Figure 4.12 (A) shows that treatment of mESCs with MLN4924 resulted in a 69% increase in Nanog protein levels, a 65% decrease in pHH3 protein levels and a 39% increase in nucleus size. Interestingly, the knockdown of CUL3 resulted in a 101% increase in Nanog protein levels but had no significant effect on pHH3, or nucleus size. Figure 4.12 (B) shows that neither cullin inhibition, nor increased Nanog protein levels had a significant effect on the levels of either total or nuclear ROS foci.

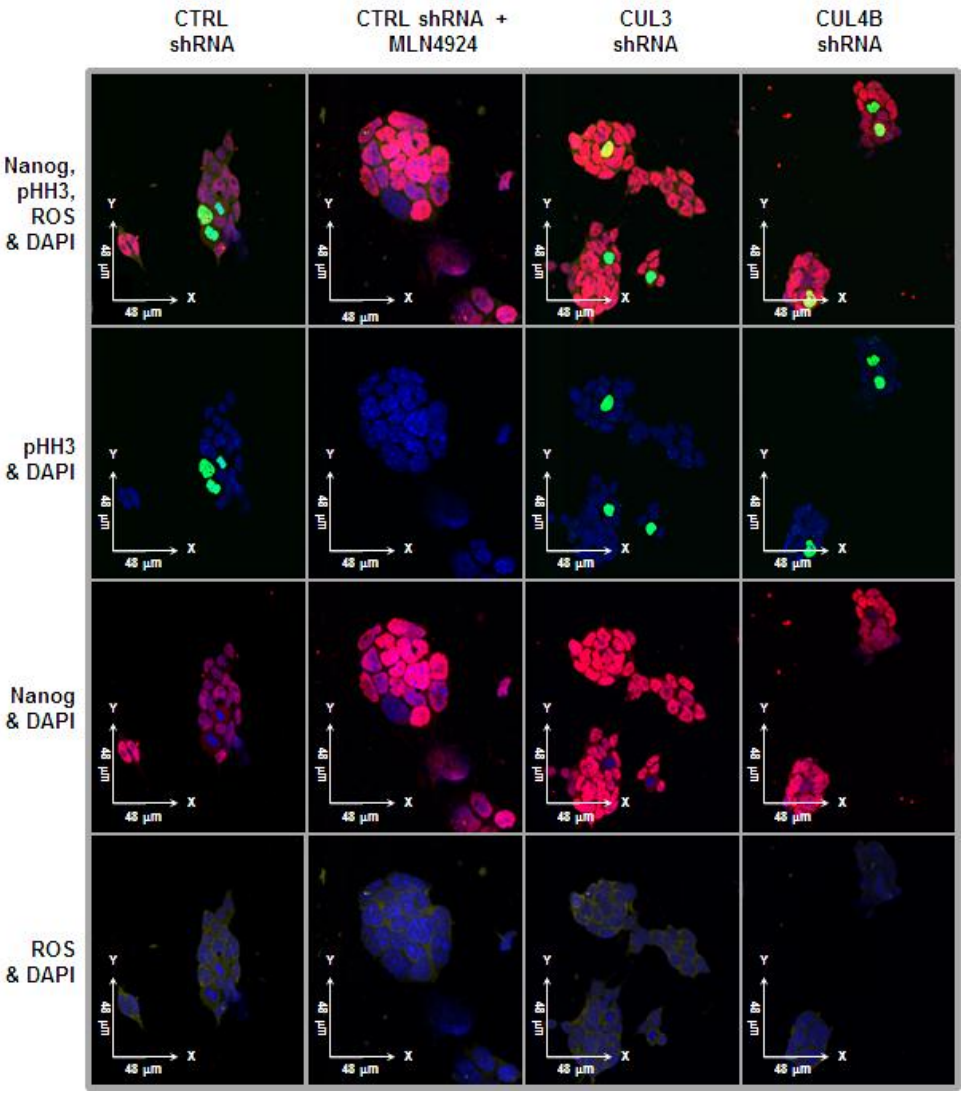
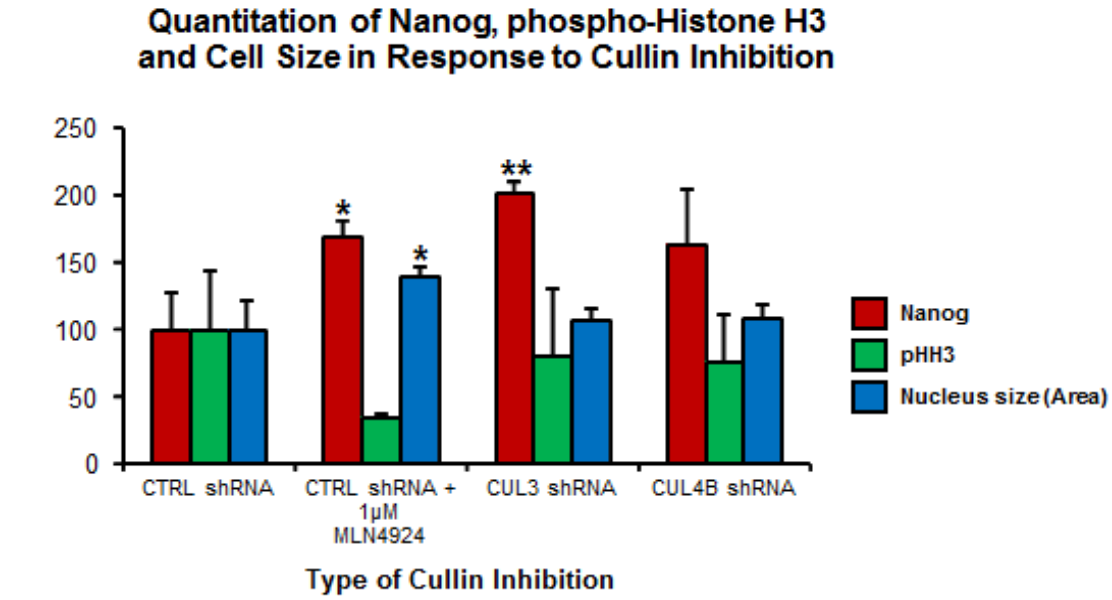


Figure 4.11: Immunostaining for Nanog, pHH3, and ROS in response to various forms of Cullin inhibition in mES cells. Overlay images of shRNA knockdown cell lines with CellROX dye for the detection of reactive oxygen species (yellow) and immunostained for Nanog (red) and pHH3 (green) proteins. Scale bar is 48 µm.

A



B

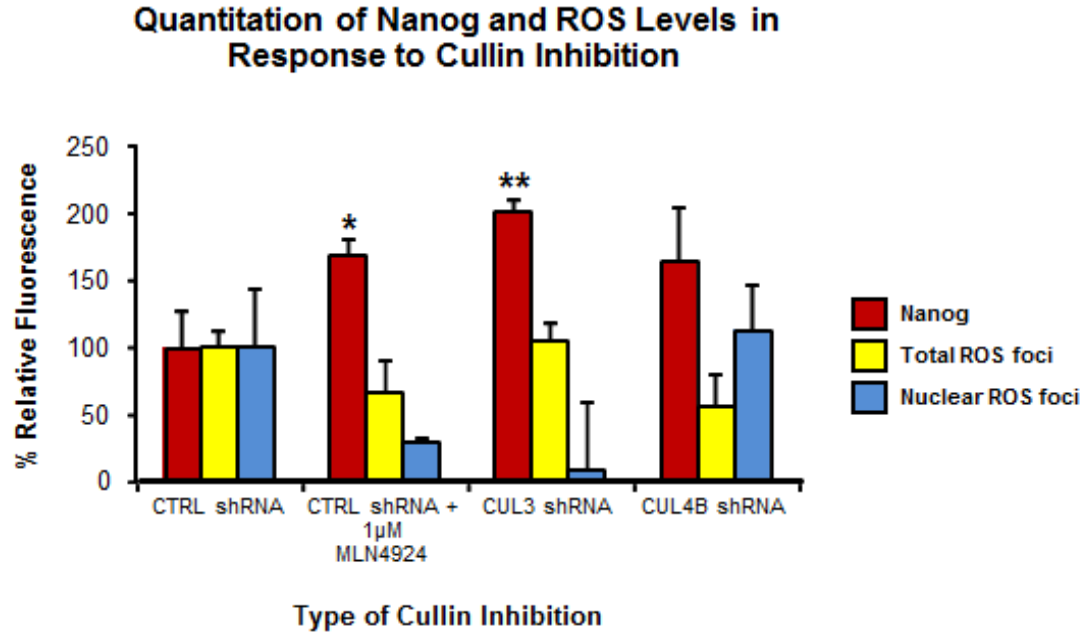


Figure 4.12: Quantitation of Nanog, pHH3, Nuclear ROS, Total ROS and cell size in response to knockdown of specific cullin proteins in ES cells. A) Nuclear Nanog protein levels relative to nucleus size and pHH3 protein levels. **B)** Nuclear Nanog protein levels relative to nuclear ROS foci and total cellular ROS foci. The % relative fluorescence was calculated using UT Control fluorescence as the denominator.1 St. Dev. above the mean % relative fluorescence is shown. The significance compared to UT control shRNA cell line was determined using a one-way ANOVA test and is indicated with asterisks above error bars. Where: * = p-value < 0.05 and ** = p-value < 0.01.

4.3.8 Quantitation of Nanog protein levels in response to CUL7 knockdown

In order to address the question of whether or not the inhibition all CRLs by the MLN4924 compound might antagonize the effects of CUL3-specific inhibition, Nanog protein levels were quantified in response to CUL7-specific knockdown in mESCs. The effects of CUL7 knockdown on Nanog protein levels were examined because it was identified in Chapter 3 of this study as part of the Nanog protein interaction network. Since both CUL3 and CUL4B knockdown resulted in an increase in Nanog levels, it was wondered if CUL7 might be responsible for the observed undesirable effects of MLN4924. That is, the moderate increases in Nanog protein levels and the enlargement of the nucleus.

Figure 4.13 (A) shows overlay images of Nanog (green) and DAPI (blue). Figure 4.13 (B) shows a 53% decrease in Nanog protein levels (within 1 St. Dev. compared to control cells) in response to CUL7-specific inhibition in mESCs. No significant changes in nucleus size (Area) were detected.

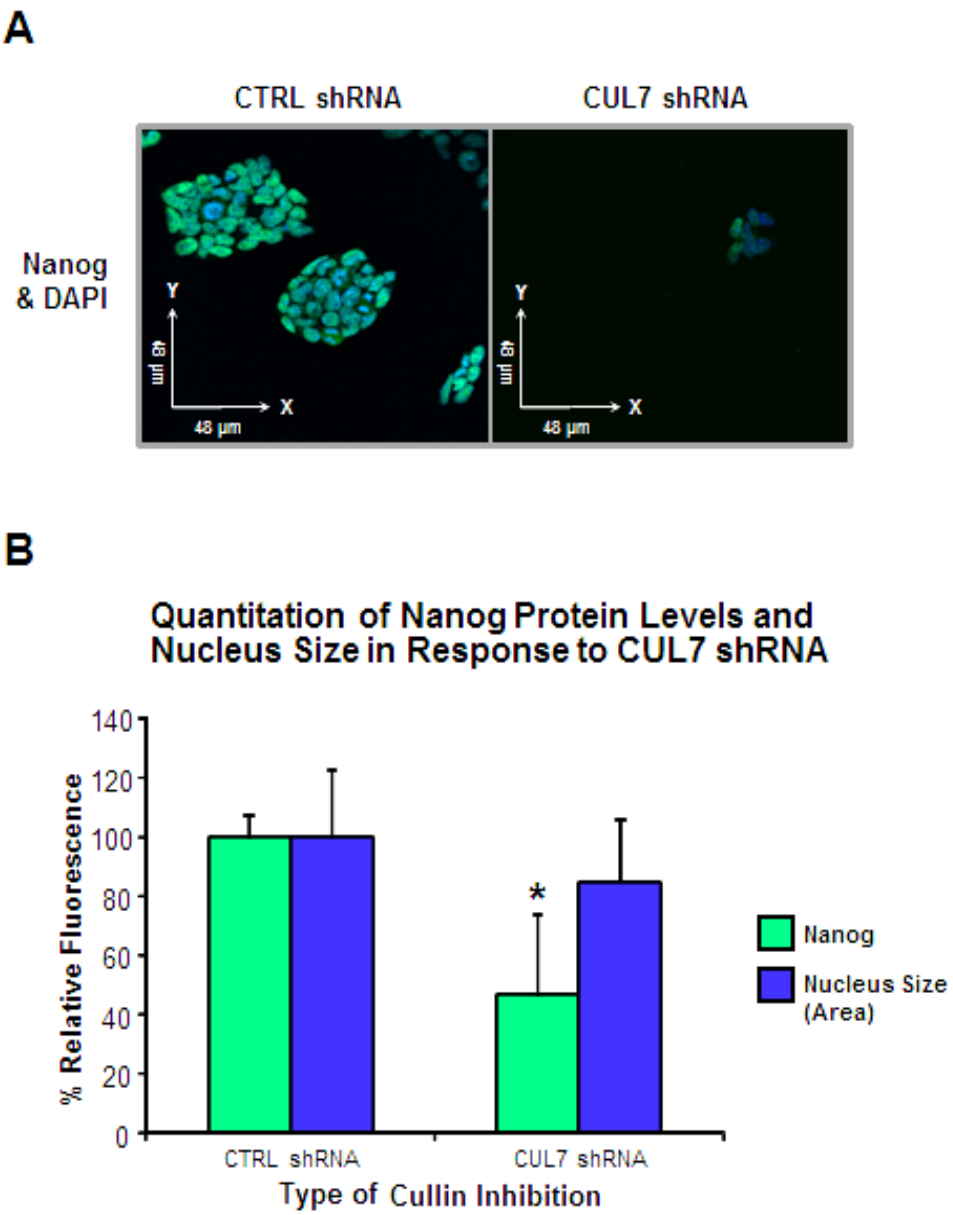


Figure 4.13: Specific knockdown of CUL7 and its effects on Nanog protein levels in ES cells. A) anti-FLAG immunofluorescent staining for Nanog protein (green) and DAPI (blue). Scale bar is 48 µm. **B)** Nanog protein levels relative to nucleus size in response to CUL7 inhibition. The % relative fluorescence was calculated using UT Control fluorescence as the denominator. 1 St. Dev. above the mean % relative fluorescence is shown. The significance compared to UT control shRNA cell line was determined using a one-way ANOVA test and is indicated with asterisks above error bars. Where: * = p-value < 0.05.

4.3.9 Quantitation of Sox2 protein levels in response to CUL4B inhibition

In Chapter 3, the CUL4B isoform was identified as an interactor with both Nanog and Sox2 proteins by IP-LC-MS/MS. In Figure 4.8 (Section 4.3.4), it was shown that the Sox2 protein and the CUL4A/B proteins co-localize. Furthermore,

this co-localization was shown to occur primarily in the cytoplasm until mitosis where this association was detected on mitotic spindles and condensed chromosomes. This suggests that the Sox2 protein is targeted by CUL4A/B-based CRL for degradation in a cell cycle-dependent manner in mES cells.

To test this hypothesis, Sox2 and CUL4A/B protein levels were quantified in response to CUL4B-specific knockdown using confocal microscopy. Data was extracted from these images using Macro3 and the mean relative fluorescence averaged from 3 z-stacks was used to calculate the percentage relative fluorescence.

Figure 4.14 (A) Shows overlay images of Sox2 (red) and CUL4A/B (green). Once again, as in Figure 4.8, the CUL4A/B protein is primarily cytoplasmic while Sox2 is distributed through both the cytoplasm and the nucleus. Conversely, in response to CUL4B-specific inhibition, the fluorescence signal for both the Sox2 protein and CUL4A/B was primarily nuclear.

Figure 4.14 (B) Shows quantitation of the Sox2 and CUL4A/B proteins, as well as, measurement of nucleus size (Area) in response to CUL4B-specific shRNA compared to control. In this experiment, Sox2 protein levels were increased by 188% in response to CUL4B knockdown. The level of immunofluorescence detected after immunostaining using an antibody recognizing both isoforms of CUL4 (CUL4A and CUL4B) proteins was increased by 286% in response to CUL4B knockdown. Additionally, the CUL4B shRNA-expressing cell line showed a 32% decrease in nucleus size (Area).

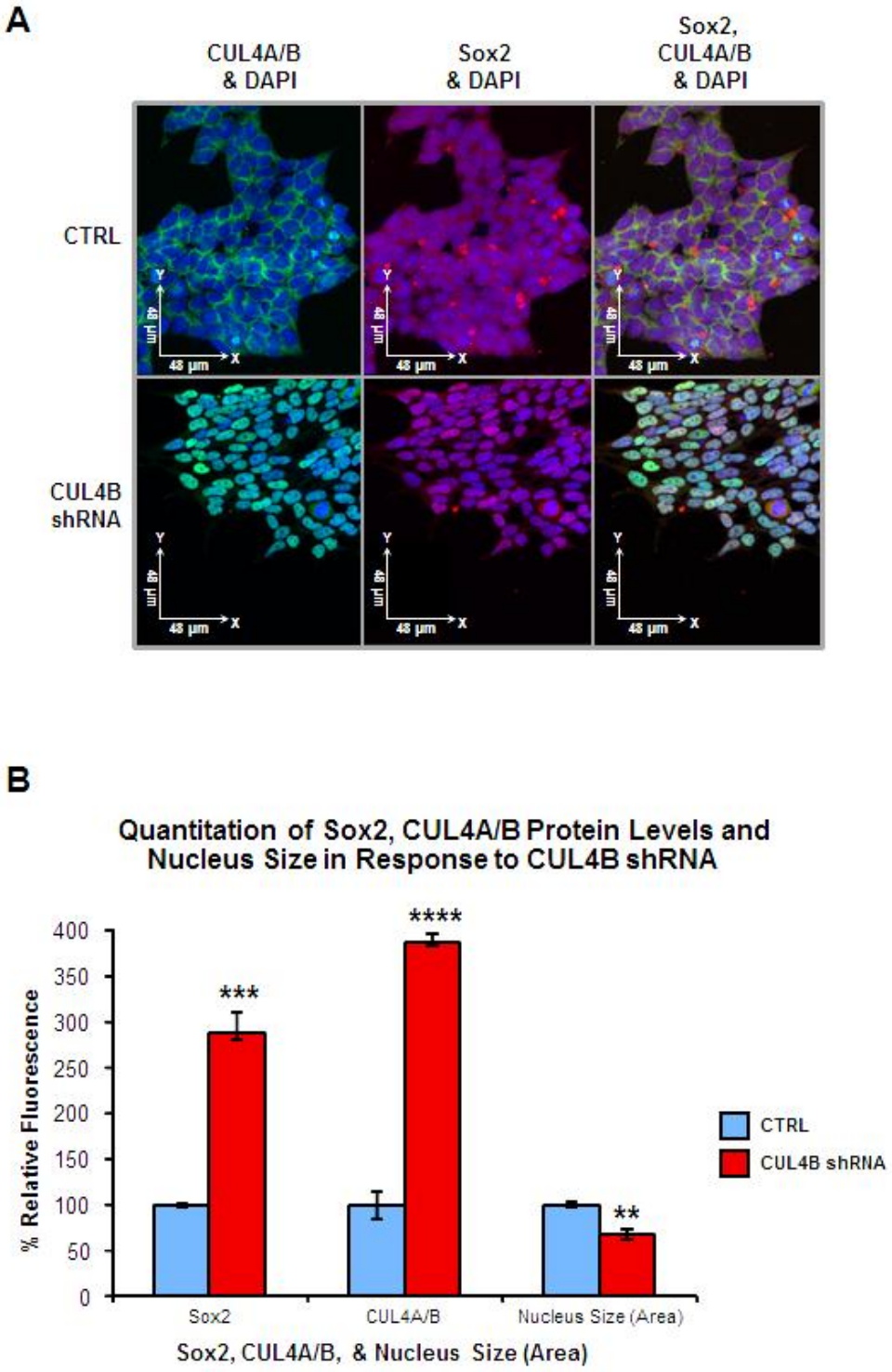


Figure 4.14: Quantitation of Sox2 protein levels in response to CUL4B knockdown in ES cells. A) Control (CTRL) shRNA and CUL4B shRNA stable cell lines immunostained with anti-CUL4A/B (green), anti-Sox2 antibody (red) and with DAPI (blue).Scale bar is 48 µm. **B)** Sox2 relative to CUL4A/B proteins levels and nucleus size in response to CUL4B inhibition. The % relative fluorescence was calculated using UT Control fluorescence as the denominator. 1 St. Dev. above the mean % relative fluorescence is shown. The significance compared to UT control cells was determined using a one-way ANOVA test and is indicated with asterisks above error bars. . Where:
* = p-value < 0.05;** = p-value < 0.01; *** = p-value < 0.001, and **** = p-value < 0.0001.

4.3.10 Co-localization of Sox2 with CUL4A/B and Nanog proteins in response to CUL4B inhibition

In Figure 4.15, a dramatic shift in the cellular localization of both the Sox2 protein and the protein being recognized by the CUL4A/B antibody (in this case, CUL4A) was observed in response to CUL4B knockdown. In this experiment, yellow co-localizing pixel overlays, as well as, Pearson's co-efficients were generated using Volocity 6.0.

Figure 4.15 (A), shows co-localization of Sox2 protein and DAPI increased in response to CUL4B knockdown. Similarly, the co-localization of CUL4A/B signal with DAPI was also increased when CUL4B-specific shRNA was expressed. The co-localization between Sox2 and CUL4A/B (CUL4B) signal was also increased in CUL4B shRNA cells. Additionally, this co-localization occurred exclusively in the nucleus in response to CUL4B shRNA. Taken together, these results indicate that Sox2 and CUL4B proteins interact in the cytoplasm where Sox2 protein levels are regulated by a CUL4B-based CRL.

Figure 4.15(B) Shows an enlarged view comparing Sox2 and CUL4A/B protein co-localization in response to CUL4B shRNA compared to untreated control cells. Red arrows point out cells at different stages in the cell cycle. In untreated control cells, Sox2 and CUL4A/B co-localization was observed in the cytoplasm until mitosis when the signals co-localized at mitotic spindles and condensed chromosomes. In response to CUL4B shRNA, however, the Sox2 and CUL4A/B signals co-localization occurred strictly in the nucleus. Upon entry into mitosis by the CUL4B shRNA-expressing cells, Sox2 and CUL4A/B signals co-localized in the nucleus, at mitotic spindles and condensed chromosomes as in control mES cells.

Figure 4.15 (C) Shows that the co-localization between the Nanog and Sox2

proteins was also increased in response to CUL4B shRNA compared to control cells.

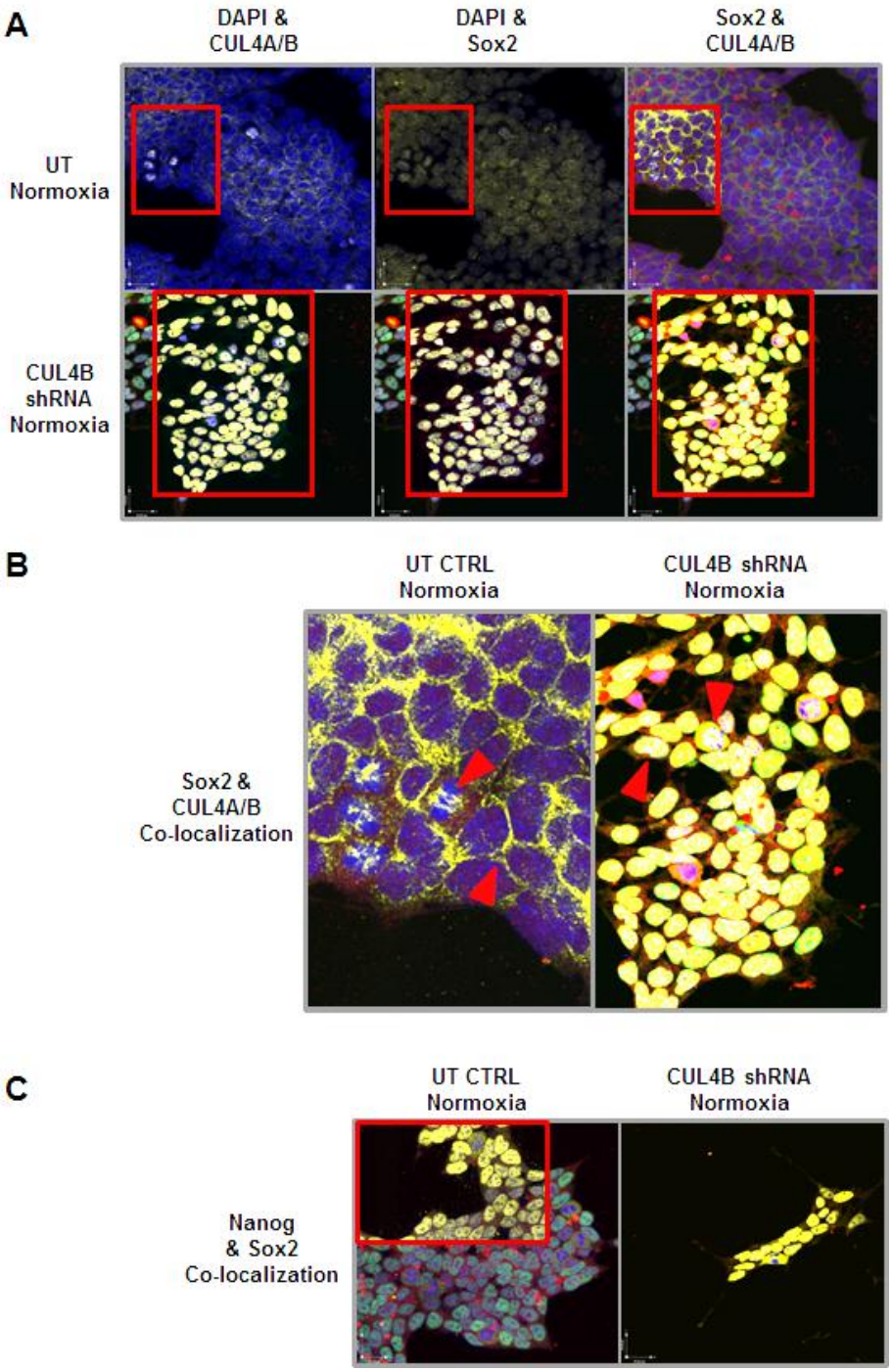


Figure 4.15: Co-localization of Sox2 with CUL4A/B and Nanog in response to CUL4B knockdown in ES cells. A) Co-localization of DAPI & CUL4A/B in response to CUL4B shRNA compared to UT control cells (Pearson's = 0.715 and 0.422, respectively). Co-localization of Sox2 protein and DAPI in these cells (Pearson's = 0.575 and 0.428, respectively). Co-localization between Sox2 protein & the CUL4A/B in response to CUL4B shRNA compared to UT control cells (Pearson's = 0.865 and 0.642, respectively). **B)** Enlarged image with red arrows indicating the cellular localization of Sox2 in association with CUL4A/B (**Left panel**) and the primary cellular localization of both the Sox2 and CUL4A/B signal is nuclear in response to CUL4B shRNA (**Right panel**). **C)** Co-localization of Nanog and Sox2 proteins in response to CUL4B shRNA (**Right panel**) compared to control cells (**Left panel**) (Pearson's = 0.856 and 0.351, respectively). Scale bar is 24 μm.

4.3.11 Cell cycle analysis of Nanog protein expression in response to the knockdown of CUL3, CUL4B and CUL7

In Figure 4.15, it was shown that Sox2 co-localized with the CUL4A/B proteins during the late stages of mitosis at mitotic spindles and condensed chromosomes. Therefore, it was investigated whether or not Nanog was also being targeted for degradation by its putative CUL3-based CRL at a specific stage of the cell cycle. In order to study Nanog protein levels throughout the cell cycle, a mitotic shake-off experiment was performed as per the protocol developed by Savatier *et al.*, 1996. As noted by Savatier, ES cells require one hour post-adhering to flasks in order to begin cycling once again and, therefore, T0 in this experiment represents 2 hours post-harvest. In Figure 4.16, western blotting was performed on whole cell lysates prepared from cells harvested at various time-points throughout this experiment. Comparisons of Nanog protein expression levels were determined relative to the mitotic cyclin, Cyclin B1.

Figure 4.16 (A) shows Nanog protein expression in CUL3 shRNA-expressing cells were detectable 8 hours earlier and was dramatically increased at 24 hours post-release into medium compared to control cells. CUL4B shRNA-expressing cells was similar in the timing Nanog expression to CUL3 knockdown. However, the same dramatic increase in Nanog protein levels at T24 in CUL3 knockdown cells was not observed in response to CUL4B knockdown. CUL7 shRNA-expressing cells had the identical Nanog expression profile as control cells. Figure 4.16 (B) shows the expression profiles of CUL3, Nanog and CyclinB1 as a function of cell cycle in parental 3XFLAG-Nanog mESCs. At T24, the levels of both CUL3 and CyclinB1 proteins drop dramatically and Nanog protein is expressed. Figure 4.16 (C) shows that CUL3 knockdown alters the expression of both CyclinB1 and Nanog. Cyclin B1 levels were overall much lower in these cells which caused problems for detection. However, the Cyclin B1 profile was similar to control cells while Nanog expression was detected 8 hours earlier than in control

cells (T16), and was dramatically increased at T24 compared to control cells. Figure 4.16 (D) shows that in response to CUL4B knockdown, an equal amount of Nanog protein was detectable at T16 and T24 post-release from nocodazole treatment.

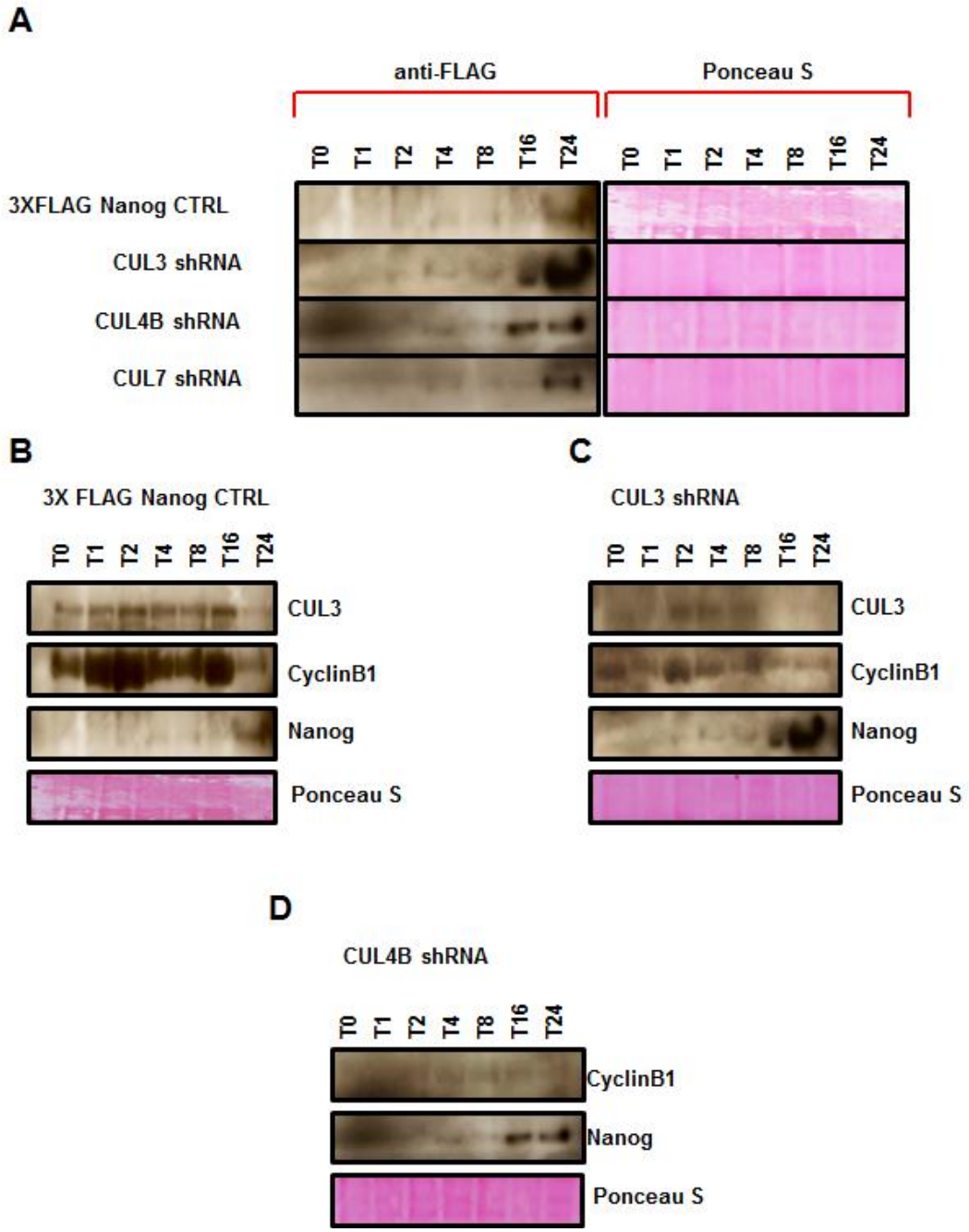


Figure 4.16: Nanog protein expression throughout the cell cycle in response to specific knockdown of CUL3, CUL4B and CUL7. A) anti-FLAG western blots Nanog protein levels post release from nocodazole synchronization in mitosis. **B)** Western blot showing the levels of Nanog protein relative to CUL3 and Cyclin B1 proteins in control cells. **C)** Western blot showing the levels of Nanog protein relative to CUL3 and Cyclin B1 proteins in response to CUL3 shRNA. **D)** Western blot showing the levels of Nanog protein relative to CUL3 and Cyclin B1 proteins in response to CUL4B shRNA. Ponceau S staining showing the amount of protein loaded per well.

4.4 Discussion:

4.4.1 Effects of various chemical compounds on the Nanog protein

As previously discussed in Chapter 3 – Section 3.4.1, Nanog is considered the 'gate-keeper' of pluripotency. It safeguards pluripotency by preventing ES cells from differentiating when Nanog levels are high and provides a window of opportunity for differentiation to occur when its levels are low (Mitsui *et al.*, 2003; Ivanova *et al.*, 2006; Chambers *et al.*, 2007). When the Nanog protein is phosphorylated, it self-dimerizes and is stabilized (Ramakrishna *et al.*, 2011; Chae *et al.*, 2012). Nanog is then able to exert it effects: increased cell proliferation, impose LIF-independence in culture and preserve pluripotency (Mullin *et al.*, 2008; Wang *et al.*, 2008). Several molecules were identified in this study as being able to increase Nanog phosphorylation above UT control cells. Namely, they are: PKI (14-22), Ht-31, MLN4924, Aurk*i*III and nocodazole. Additionally, a few others were identified which cause the disappearance of the phospho-Nanog protein band: H89 and Wortmannin. These results revealed for the first time to my knowledge, a method of controlling phospho-Nanog protein levels.

4.4.2 Effects of various perturbations on Nanog-containing protein complexes

In this section, native gel analysis of Nanog protein complexes purified from cells treated with H89 demonstrated that the disappearance of the phospho-Nanog protein band (Figure 4.1) de-stabilizes Nanog complexes in mESCs. Further analysis by 2D Native-to-SDS-PAGE confirmed the de-stabilization of Nanog complexes by this compound. Although the results are quite interesting and H89 is advertised and sold as a PKA inhibitor, H89 has well documented off-target effects. In fact, this

compound has been shown to inhibit 8 other kinases. They are: MAPKAP-K1b, MSK1, PKBα, SGK, S6K1, ROCK II, AMPK, and CHK1 (Davies *et al.*, 2000; Lochner *et al.*, 2006). For this reason, a more specific method of inhibiting PKA was required in order order to confirm the results obtained using H89. One specific method of PKA inhibition is by the cell permeable peptide sequence (14-22) derived from heat-stable protein kinase inhibitor (PKI). PKI is an endogenous inhibitor of PKA and peptide derivatives thereof have been shown to effectively and specifically inhibit PKA activity both *in vitro* and *in vivo* (Glass *et al.*, 1989; Walsh *et al.*, 1991). Therefore, a commercially available cell-permeable version of this peptide was tested alongside H89. As shown in Figure 4.1, treatment of 3XFLAG-Nanog mESCs with PKI (14-22) did not confirm the results obtained with H89. Conversely, the specific inhibition of PKA with this molecule greatly increased phospho-Nanog levels. Therefore, the depletion of phospho-Nanog by H89 treatment appears not to be due to the compound's reported activity as a PKA inhibitor but rather by inhibiting one or more of its other 'off-target' kinases.

In addition to the fact that H89 inhibits 8 other kinases, Davies *et al.*, 2000 showed that three kinases MSK1, S6K1 and ROCKII are inhibited with an IC50 of 120 nM, 80 nM, 270 nM, respectively. These values are close to or below that reported for H89 inhibition of PKA, 135 nM. Using chemical compounds or small molecules that specifically inhibit each of these kinases, the depletion of phospho-Nanog protein levels should be further investigated.

4.4.2.1 Model: The Molecular Keys to the Gateway of Pluripotency

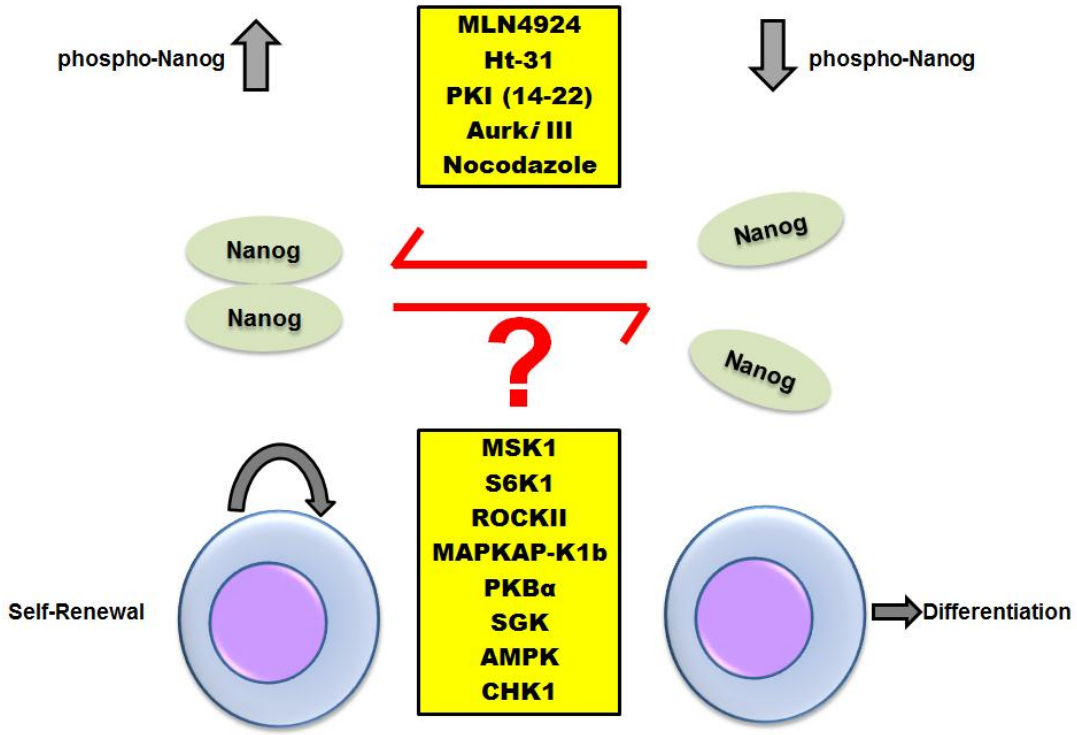


Figure 4.17: The molecular keys that lock and unlock the gateway to pluripotency. The chemical compounds and small molecules shown in this study to increase or decrease phospho-Nanog levels in ES cells. Where increased phospho-Nanog induces self-dimerization and stabilizes Nanog function: maintains the pluripotent state. Conversely, decreased phospho-Nanog levels prevents dimer formation, pluripotency is lost and cells are susceptible to differentiation.

Figure 4.17 shows that Nanog dimerization mediated by phospho-Nanog can be either stabilized or inhibited using small molecules. In the yellow boxes are the compounds identified in this study to modulate phospho-Nanog levels. The red arrows indicate the direction in which the compounds listed in the yellow boxes are driving Nanog phosphorylation and stabilization. The red question mark above the list of kinases driving the reaction toward de-stabilization, indicates that it is currently not known which one or combination of these kinases is responsible for the multi-site phosphorylation of Nanog. Therefore, inhibitors to each of these kinases

both individually and in combination should be tested for their effects on phospho-Nanog levels.

4.4.3 Quantitation of Nanog, Sox2 and CUL4A/B proteins in response to inhibition of the PKA signaling pathway

Since two of the molecules shown to increase phospho-Nanog levels in this study inhibit PKA's ability to phosphorylate its substrates, these compounds were further tested for the effects on the levels of Nanog, Sox2 and CUL4A/B proteins. In this section an attempt was made to further dissect the PKA signaling pathway by quantitating both the total cellular, as well as, the nuclear levels of these 3 proteins.

The specific inhibitor of PKA activity, PKI (14-22) was employed as well as the specific inhibitor of AKAP signaling, Ht-31 (Carr *et al.*, 1991; Wang *et al.*, 2006). Forskolin, an activator of adenylyl cyclase was employed alongside a more specific method of elevating intracellular cAMP levels using cell-permeable Br-cAMP. This comparison was made in order to address whether the disappearance of phospho-Nanog caused by Forskolin (Figure 4.1) was due to its ability to elevate cAMP levels by activating adenylyl cyclase (Seamon *et al.*, 1981; Insel *et al.*, 2003). All 4 molecules were compared to both UT normoxia control cells as well as UT cells cultured in hypoxia. The comparison with hypoxia was done in order to determine whether or not the use of small molecules to increase phospho-Nanog levels was able to recapitulate physiologically relevant growth conditions in ES cells.

By extracting the data using two different macros (Macro1 and Macro2), it was possible to quantify the levels of Nanog, Sox2 and CUL4B proteins throughout the entire cell as well as specifically in the nucleus. Using Macro1, no significant changes in the total levels of these proteins was detected in response to any of the

small molecule and peptide treatments employed. In hypoxic conditions, however, the total levels of all three proteins were significantly increased (Figure 4.6).

The quantitation of nuclear Nanog, Sox2 and CUL4A/B protein levels revealed that specific PKA pathway inhibition elevated Sox2 levels in mES cells (Figure 4.7). However, hypoxia-cultured cells showed the greatest increase in nuclear Sox2 levels compared to any of the treatments tested. Interestingly, the cell permeable peptide inhibitors PKI and Ht-31 (potent elevators of phospho-Nanog levels) increased levels of nuclear Sox2 protein significantly above controls. None of the treatments tested had any significant effect on nuclear Nanog protein levels. Therefore, an increase in phospho-Nanog levels does not increase overall Nanog protein levels. Taken together with the results shown in Figure 4.1, this data indicates that PKI (14-22) and Ht-31 do not increase overall Nanog protein levels but, instead, induce a possible equilibrium between the phosphorylated and unphosphorylated forms.

Considering, the significant increase in nuclear Sox2 levels combined with the fact that Ht-31 treatment also enriches for the formation of Nanog-Sox2 protein complexes suggests that both the phosphorylation of Nanog, it predicted dimerization (Ramakrishna *et al.*, 2011; Chae *et al.*, 2012) as well as the Nanog-Sox2 interaction are necessary to prevent differentiation.

4.4.3.1 Model: The molecular keys that unlock the gateway to pluripotency

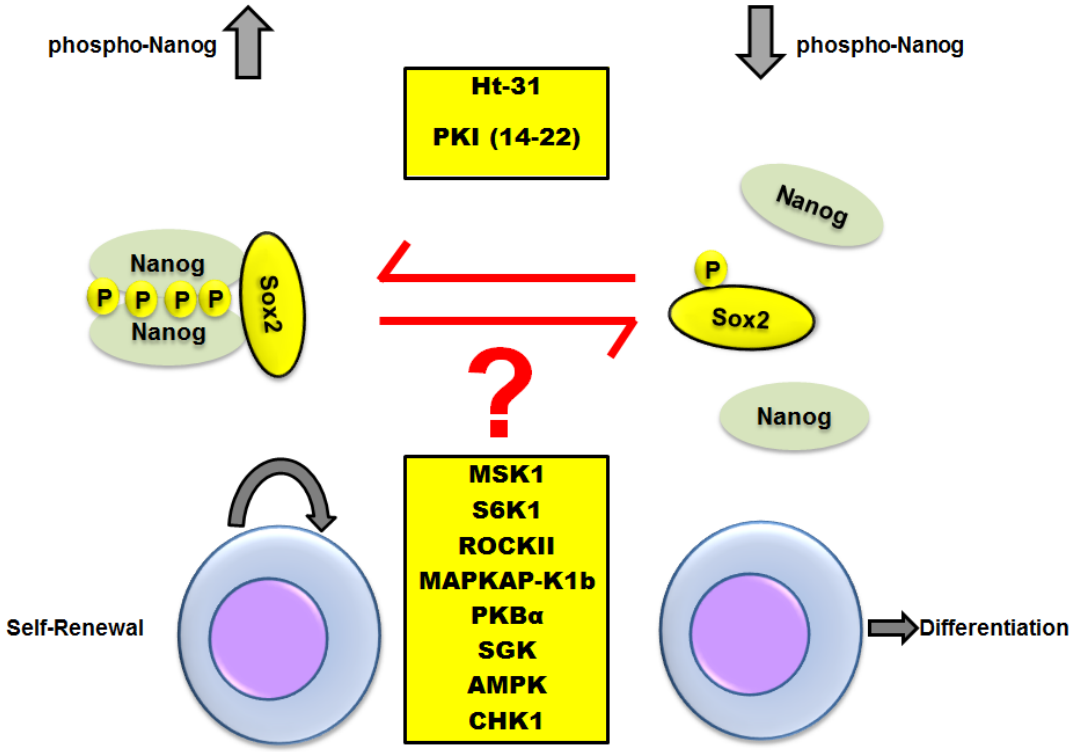


Figure 4.18: The molecular keys that lock and unlock the gateway to pluripotency. In this model, the PKI (14-22) and Ht-31 prevent PKA from phosphorylating Sox2 thereby stabilizing the Sox2 protein and its interaction with the phospho-Nanog/Nanog dimer. The interaction of Sox2 with the dimerized form of Nanog is required to prevent the disappearance of phospho-Nanog possibly via ubiquitination and targeting it for degradation. When the Nanog dimer and its interaction with Sox2 are stabilized, the doorway leading to differentiation is closed.

This model is based on the data presented thus far in this study. Although Sox2 has never been reported to be phosphorylated by PKA, it has been reported to be phosphorylated by PKB/Akt1 on Threonine 118 (Jeong *et al.*, 2010). Comparison of the phosphorylation motifs recognized by PKA and PKB/Akt1 reveals that the two recognition sequences are highly similar (Table 4.3).

Kinase	Motif
PKA	RRX(S/T)*
PKB	RXX(S/T)*

Table 4.3 The substrate recognition sequences of PKA and PKB/Akt1. Phospho-motif sequences from Lawlor *et al.*, 2001. * Indicates the amino acid can be either Serine or Threonine.

In this particular case, the site identified by Jeong *et al.*, 2010 to be phosphorylated by PKB/Akt1 can also be phosphorylated by PKA. In this study, the PKA-anchoring proteins Akap8 and Akap8l, were both identified as interactors with Nanog by IP-LC-MS/MS, as well as, Sox2. Furthermore, the amino acid that Jeong *et al*, 2010 reported as being phosphorylated on Sox2 (Threonine 118) lies within the High Mobility group (HMG) domain which is conserved in all Sox proteins. Previous studies on Sox9 have shown that this protein is phosphorylated within its HMG domain at Serine 181 by PKA (Huang *et al.*, 2000; Malki *et al.*, 2005). Alignment of the Sox9 and Sox2 HMG domains reveals that Serine 181 in the Sox9 protein corresponds to Threonine 118 in the Sox2 protein (Figure 4.19).

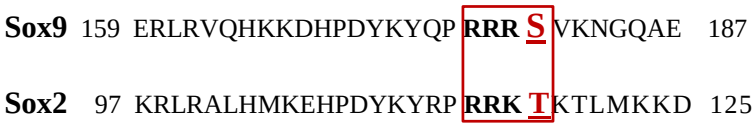


Figure 4.19: The PKA and PKB/Akt1 phospho-motif within the conserved HMG domain of Sox9 and Sox2. The PKA recognition sequence identified in Sox9 is also found in the Sox2 protein sequence. Red box indicates the PKA phospho-motif found in both proteins. Red letters indicate the residues being phosphorylated by PKA.

It has also been reported that PKA inhibition prevents the phosphorylation of Oct4 and further activates Oct4 mediated transcription (Saxe *et al.*, 2009). Therefore, the observed increases in nuclear Sox2 protein levels could also be attributed, at least in part, to increased transcription of the Sox2 gene by Oct4.

4.4.4 Co-localization of Nanog, Sox2 and CUL4A/B

Introduced in Chapter 3 was the decision to perform the IP-LC-MS/MS study using whole cell lysates rather than nuclear extracts. This decision was partly based on the documented life of Sox2 outside the nucleus (Li *et al.*, 2007; Baltus *et al.*, 2009).

In this study CUL4B was identified as an interactor with both Sox2 and Nanog by IP-LC-MS/MS. This interaction was also observed by Pardo *et al.*, 2010 in their interaction network for Oct4. Since Pardo *et al*., 2010 also performed their IP-LC-MS/MS study on whole cell lysates, it was hypothesized that this interaction was occurring in the cytoplasm rather than the nucleus.

In order to test this hypothesis, the 3XFLAG-Nanog cell line was immunostained for Nanog, Sox2 and CUL4A/B. Co-localization of these proteins was determined by confocal microscopy. In Figure 4.8, it was shown that Sox2 and Nanog co-localization occurred in the nucleus. This suggests that a life beyond the nucleus is most likely restricted to the Sox2 protein.

Sox2 and CUL4A/B co-localization was detected in the cytoplasm until mitosis when their co-localization was observed at mitotic spindles and surrounding condensed chromosomes. This observation provides insight into why Nanog was not reciprocated in Sox2 IP-LC-MS/MS experiments.

In the case of Nanog and CUL4A/B, co-localization occurred solely in the nucleus. As was the case for Sox2, this co-localization was also observed at mitotic spindles and surrounding the condensed chromosomes. Furthermore, CUL4A/B co-localized with DAPI very weakly until mitosis when it was detected on condensed chromosomes.

It is important to point out that due to the lack of a CUL4B-specific antibody, the antibody used in this study recognizes both isoforms of CUL4; CUL4A and CUL4B. Thereby, making it impossible to distinguish which of the one or both isoforms are co-localizing with Nanog and Sox2 in control cells when both isoforms are being expressed.

From the results in this experiment it would appear as though the hypothesis is correct: the identification of CUL4B by IP-LC-MS/MS in this study and the Pardo *et al.*, 2010 was due to whole cell lysates being used as opposed to nuclear extracts and this interaction occurs in the cytoplasm. However, this is counter-intuitive since CUL4B has been reported to be a nuclear protein whereas CUL4A is both nuclear and cytoplasmic (Guerrero-Santoro *et al.*, 2007; Zou *et al.*, 2009). The immunostaining in this experiment shows that the majority of the CUL4A/B protein is cytoplasmic with very little detectable nuclear fluorescence. That is, until mitosis where it translocates to the nucleus to condensed chromosomes and mitotic spindles. This cell cycle-dependent localization is characteristic of CUL4 and has been previously shown for both CUL4A and CUL4B isoforms (Higa *et al.*, 2006; Jin *et al.*, 2006; Guerrero-Santoro *et al.*, 2008).

Despite this, only peptides specific to the CUL4B sequence were detected by LC-MS/MS in the Nanog and Sox2 IPs. Therefore, in order to confirm our hypothesis, further investigation into whether or not the primarily cytoplasmic staining is indeed due to CUL4B was done (see Sections 4.4.10 and 4.4.11). The knockdown of CUL4B expression levels in mESCs resulted in the almost complete disappearance of the CUL4A/B antibody signal in the cytoplasm. Therefore, this confirms that CUL4B is expressed most abundantly in the cytoplasm of mESCs

stably expressing high levels of the Nanog protein. In contrast, CUL4A is predominantly localized in the nucleus of these cells.

4.4.5 Effects of the chemical compound MLN4924 on the protein levels of various pluripotency factors

In Chapter 3, three cullin proteins were identified as interacting with pluripotency factors by IP-LC-MS/MS. In Section 4.3.1, a compound that inhibits all Cullin-RING Ligase (CRL) activity in cells, the NAE1 inhibitor MLN4924, was identified as one of several compounds able to stabilize phospho-Nanog protein levels. Since CRL substrates are typically phosphorylated prior to being ubiquitinated by their respective CRLs (Hershko *et al.*, 1998; Glickman *et. al.*, 2002; Lu *et al.*, 2009), the observed increase in phospho-Nanog levels following treatment with MLN4924 supports the notion of CRL-mediated degradation of Nanog.

In addition to Nanog, Sox2 and Esrrβ were also found to interact with cullin proteins. Both CUL4B and CUL3 were found to interact with Sox2 and CUL7 was detected in the Esrrβ purifications (Chapter 3). In order to determine if the Sox2 and Esrrβ proteins might also be substrates for CRL-mediated degradation, their levels were examined in response to MLN4924 treatment.

The levels of phospho-Nanog were increased following treatment with MLN4924 for 12 and 24 hours. Esrrβ protein levels were unaffected after 12 hours and 24 hours of treatment with the compound. However, a significant effect on Sox2 protein levels was observed. The levels of Sox2 protein increased following 12 hour treatment with MLN4924 and were further increased after 24 hours. At the same time, CUL4A/B protein levels also increased over time. This effect can be attributed to the MLN4924. Since this compound has been shown to specifically inhibit the

NAE1 enzyme in both mammals and plants (Soucy *et al.*, 2009; Milhollen *et al.*, 2009; Hakenjos *et al.*, 2007), the increase in CUL4A/B protein levels can be explained by their documented ability to undergo auto-ubiquitination (Groisman *et al.*, 2003; Hrecka *et al.*, 2007; Jackson *et al.*, 2009). The inhibition of CUL4-based CRLs activity would, therefore, prevent both CUL4 isoforms from self-destruction and stabilize CUL4A/B protein levels as well as their substrates.

4.4.6 Effects of individual cullin-specific shRNA on the interaction of Nanog with Sox2 and Esrrβ

The purification of the Nanog protein from ES cells stably expressing CUL3, CUL4B and CUL7 shRNA was performed in order to determine whether or not its interaction with the Sox2 and Esrrβ proteins was due to Nanog being degraded by the same CRL as these proteins (Section 4.3.6).

In this experiment, the amount of Nanog protein being purified from ES cells was increased when CUL3 or CUL4B were knocked down. This implies an overall increase in the amount of Nanog protein in ES cells when these genes were knocked known. The greatest increase in the amount of Nanog protein being purified occurred when CUL3 was knocked down. This increase was significant compared to cells stably expressing a control scrambled shRNA construct and strongly suggests that CUL3 may be responsible for the degradation of Nanog .

Esrrβ was found to co-purify with Nanog at the same levels regardless of which cullin was knocked down. This result indicates that the Esrrβ interaction with Nanog is not due to their being degraded by the same ligase and therefore Esrrβ can be considered a true interaction with the Nanog network.

In the case of Sox2, an increasing amount of protein was observed to co-

purify with Nanog in response to CUL3, CUL4B and CUL7 shRNA, respectively. This result indicates that Sox2 is an interactor with Nanog and, this interaction is not the result of both proteins co-localizing with (as demonstrated in Section 4.3.4) and being degraded by CUL4B.

4.4.7 Quantitation of Nanog and phospho-Histone H3 protein levels in response to cullin inhibition

It was demonstrated in Section 4.3.7 that treatment with MLN4924 increases the levels of a phosphorylated form of the Nanog protein, however, total levels of Nanog appeared relatively equal. In this experiment, Nanog protein levels were quantified and revealed that MLN4924 increases Nanog protein levels by 30%. Therefore, the increase in phospho-Nanog due to CRL inhibition does correspond to an increase in total Nanog protein levels within the cells. Furthermore, this indicates that there may be more than one phospho-Nanog protein exists in mESCs: one which stabilizes Nanog dimerization and pluripotency and another possibly differentially phosphorylated form of Nanog which is being targeted for ubiquitination by a CRL.

In Section 4.3.6, the amount of Nanog protein purified from ES cells in response to CUL3 and CUL4B inhibition was increased. The results from this experiment indicate that Nanog protein levels are further increased by the specific inhibition of CUL3 by shRNA. This suggests that Nanog is being degraded by a CUL3-based CRL. In response to CUL4B-specific knockdown, the mean percentage relative fluorescence detected for Nanog was increased to a similar level as with MLN4924. However, this result was not significant within 1 Std. Dev. of the mean compared to CTRL cells. This result implies that although Nanog levels are

increased in response to CUL4B knockdown, this effect is not due to CUL4B targeting Nanog for degradation.

A significant decrease in phospho-Histone H3 levels was observed in response to MLN4924. This implies that when ES cells are treated with MLN4924 they may be prevented from entering mitosis. Another observed effect of MLN4924 on ES cells was a significant increase in nucleus diameter by 10% above control. Although this increase may sound slight, the observed effects of cell nucleus morphology were not. Figure 4.11 shows the significantly enlarged nuclei of ES cells in response to MLN4924 treatment.

The effects of MLN4924 observed here can be explained by recently published studies. Firstly, this compound prevents CUL4-based degradation of Cdt1 in S-phase. This results in DNA re-replication, checkpoint activation and apoptosis in cancer cells (Lin *et al.*, 2010; Milhollen *et al.*, 2011). This explains the decrease in phospho-HH3 levels observed in this study as cells would be arrested in S-phase and prevented from entering mitosis. Furthermore, this also explains the observed increase in nucleus size as DNA re-replication and apoptosis would both contribute to the swelling of the nucleus. Additionally, MLN4924 has been shown to stabilize the cell cycle inhibitor p21and inhibit proliferation in tumour cells (Jia *et al.*, 2011).

Although the levels Nanog are increased by MLN4924 treatment it cannot overcome one or more of the above-mentioned effects. Altogether, the toxic effects that have been reported in the literature for MLN4924 with rapidly proliferating cancer cells can explain why the compound does not increase Nanog levels as greatly as seen for CUL3-specific inhibition. They also explain the adverse effects that

MLN4924 treatment has on ES cells, the swelling of the nucleus and preventing entry into mitosis.

4.4.8 Quantitation of Nanog protein levels in response to CUL7 knockdown

Another cullin protein which was identified as an interactor with Nanog and Esrrβ (Chapter 3) was CUL7. In order to further confirm that a CUL3 CRL is degrading Nanog, it was required to quantify the effects of CUL7 inhibition on Nanog protein levels. Additionally, it was postulated in Section 4.3.7: Quantitation of Nanog and phospho-Histone H3 protein levels in response to cullin inhibition, that although MLN4924 increases Nanog protein levels, its inhibition of all CRLs may be working in opposition to its inhibition of the CRL responsible for degrading Nanog. This is confirmed by the recent studies (Lin *et al.*, 2010; Milhollen *et al.*, 2011; Jia *et al.*, 2011), with regards to the inhibition of CUL4 CRLs. However, the role of CUL7 CRL-specific inhibition in antagonizing the increase in Nanog levels in ES cells has not been determined.

In this experiment, the levels of Nanog protein were decreased by 18% compared to CTRL shRNA-expressing cells. This effect can be attributed to the accumulation of p53, the well-known substrate of CUL7 CRLs (Andrews *et al.*, 2006; Kasper *et al.*, 2006; Kim *et al.*, 2007; Sarikas *et al.*, 2008). The repression of Nanog's expression by p53 is well-documented in the literature (Lin *et al.*, 2005; Solozobova *et al.*, 2009; Zbinden *et al.*, 2010; Brandner *et al.*, 2010; Kuijk *et al.*, 2010; Moon *et al.*, 2011). These experiments suggest that the increased Nanog levels by MLN4924's inhibition of CUL3 CRL activity is being counter-balanced at the

transcriptional level by the accumulation of p53 which results from inhibiting CUL7 CRL function.

4.4.9 Quantitation of Sox2 protein levels in response to CUL4B inhibition

There were two questions to be answered by performing this experiment: 1) Is Sox2 a substrate for CUL4B isoform-specific CRL degradation? and 2) Is the primarily cytoplasmic protein being recognized by the CUL4A/B antibody used in this study CUL4B or CUL4A?

The first question was answered by the quantitation of Sox2 protein levels in response to CUL4B-specific inhibition using shRNA. In this experiment, we see that Sox2 protein levels were increased by 163% above control cells. Also, the levels of CUL4A protein (recognized by the CUL4A/B antibody used) were increased by 262% above controls. In addition, the localization of the CUL4A/B signal was primarily nuclear as opposed to cytoplasmic as seen in the control. This increase in Sox2 and CUL4A proteins coincided with a 25% increase in nucleus size.

These results indicate that Sox2 is indeed the substrate of a CUL4B isoform-specific based CRL. In addition, this interaction was identified by IP-LC-MS/MS in Chapter 3 of this study due to the use of whole cell lysates as the starting material for the purifications performed. The identification of CUL4B in Nanog IPs is, therefore, the likely to be the result of Nanog's interaction with Sox2 observed in this study. Although Nanog staining is primarily nuclear, all 3 proteins do co-localize on mitotic spindles and condensed chromosomes during mitosis as shown, to my knowledge for the first time in this study (Figure 4.8).

The fact that this degradation occurs primarily in the cytoplasm of ES cells, reveals that CUL4B is differentially localized in ES cells. It cannot be ruled out that this altered localization might be a consequence of the ectopic expression of Nanog in these cells which, as a result, stably express high levels of the Nanog protein.

The increase in CUL4A signal (as indicated by the CUL4A/B antibody in CUL4B shRNA-expressing cells) can be attributed to an increase in the expression of the CUL4A isoform in response to the stable knockdown of CUL4B. Although the CUL4A and CUL4B isoforms have always been thought to be functionally redundant (Higa *et. al.*, 2006a; Jin *et al.*, 2006; Higa *et al.*, 2006b; Olma *et al.*, 2009), recent studies have shown that CUL4B isoform-specific CRLs have been characterized (Nakagawa *et al.*, 2011; Li *et al.*, 2011; Green and Gozani, 2011).

The identification in this study of a CUL4B isoform-specific CRL degrading Sox2 in embryonic stem cells is both novel and intriguing. As previously mentioned, these two isoforms have always been regarded as having redundant function in cells. However, they are isoforms which are encoded by separate genes with the CUL4B gene being encoded on the X-chromosome (Kerzendorfer *et al.*, 2007). Mutations in the CUL4B gene cause X-linked mental retardation and an increased sensitivity to double stranded DNA breaks (Zou *et al.*, 2007; Kerzendorfer *et al.*, 2010). Recently, a CUL4B isoform-specific CRL targeting Wdr5 for degradation was reported (Nakagawa *et al.*, 2011). Wdr5 is a component of the H3K4 methyltransferase complex involved in epigenetic repression of transcription and was also identified of the interaction network (Chapter 3 of this study). Taken together, with the CUL4B isoform-specific CRL identified here to degrade Sox2, an argument could be made that in ES cells CUL4B is the preferentially expressed isoform of CUL4. This

suggests that there are potentially several novel and currently uncharacterized CUL4B isoform-specific CRLs that are both specific and unique to ES cells.

4.4.10 Co-localization of Sox2, Nanog and CUL4A/B protein levels in response to CUL4B inhibition

This experiment revealed that in the 3X FLAG Nanog cell line CUL4B staining is primarily cytoplasmic with the exception of mitosis where once again, both proteins co-localized on mitotic spindles and on condensed chromosomes. In response to CUL4B-specific inhibition, Sox2 staining is entirely nuclear with only faint staining in the cytoplasm. The elevated CUL4A/B signal (now mostly attributed to CUL4A) is also primarily nuclear. Once again, however, both CUL4A and Sox2 proteins co-localize at mitotic spindles and condensed chromosomes. These results imply that Sox2 does indeed co-localize with the CUL4B isoform in the cytoplasm of ES cells but may associate with either CUL4A or CUL4B at condensed chromosomes and on mitotic spindles. This would explain why cells expressing only the CUL4A isoform are still able to progress through mitosis and do not undergo cell cycle arrest. This study illustrates that the CUL4A and CUL4B isoforms are not entirely functionally redundant as once thought. Where redundancy was found in this study was at mitotic spindles and condensed chromosomes. During mitosis it appears as though, in the absence of CUL4B, the CUL4A protein is acting upon their shared substrates in order to allow for cell cycle progression.

It is now becoming accepted that these isoforms have unique and specific functions which, in the case of CUL4B, involves the regulation of ES cell-expressed proteins which include: Wdr5 (Nakagawa *et al.*, 2011), Peroxiredoxin III (Li *et al.*, 2011), DNA damage repair proteins (Kerzendorfer *et al.*, 2010), chromatin

remodeling (Green and Gozani, 2011) and the master pluripotency regulator, Sox2 (to my knowledge only identified in this study).

4.4.11 Cell cycle analysis of Nanog protein expression in response to the knockdown of CUL3, CUL4B and CUL7

Mitotic shake-off experiments revealed differences the cell cycle-related expression profiles obtained for Nanog in response to CUL3 and CUL4B knockdown. Post-release into regular growth medium, the expression of Nanog begins at T24 in control cells. In response to CUL3 knockdown, Nanog expression began at T16 which is 8 hours earlier than control cells. Furthermore, at T24, a dramatic increase in Nanog levels was seen in CUL3 shRNA-expressing cells. In response to CUL4B knockdown, Nanog expression also began at T16 and persisted at T24. In contrast to CUL3 knockdown cells, however, a dramatic increase in Nanog levels did not occur at T24 in the CUL4B knockdown cells. Instead, Nanog levels were equivalent at T16 and T24 post-release. CUL7 knockdown cells had the same timing of Nanog expression as control cells at T24. Taken together, these results further confirm the hypothesis that Nanog is degraded by a CUL3 CRL.

In control cells we saw that Cyclin B1 levels began to accumulate atT4 post-release and peaked at T16. This indicates that cells began to enter mitosis 4 hours after being released into growth medium. The peak in Cyclin B1 levels, indicating the cells had entered metaphase, occurred at T16 post release. A dramatic decrease in Cyclin B1 levels was seen at T24 which indicates that cells had exited mitosis. At this same time point, Nanog protein was detected for the first time post-release.

CUL3 knockdown resulted in Nanog expression being detectable in low

amounts at T16 post-release and were dramatically increased at T24. Due to Nanog's documented role in promoting the G1/S transition in ES cells, (Zhang *et al.*, 2009; Ma *et al.*, 2009). This suggests that the cells began to enter S phase 8 hours earlier than in control cells.

Unfortunately, the detection of Cyclin B1 in the CUL3 and CUL4B experiments was so weak that no observations could be made confidently about the effects this altered Nanog expression profile may have on CyclinB1 level in these cells.

Although others have reported Nanog to be ubiquitinated and degraded by the 26S proteasome (Moretto-Zita *et al.*, 2010; Ramakrishna *et al.*, 2011), this study is (to my knowledge) the first to identify that Nanog is specifically degraded by a CUL3 CRL in ES cells. Taken together, the results obtained in Sections 4.3.4 – 4.3.12 of this study combine to form to the models being proposed in the following section: Section 4.4.11.

4.4.11 Model of Cell Cycle-dependent CUL3 CRL-mediated degradation of Nanog:

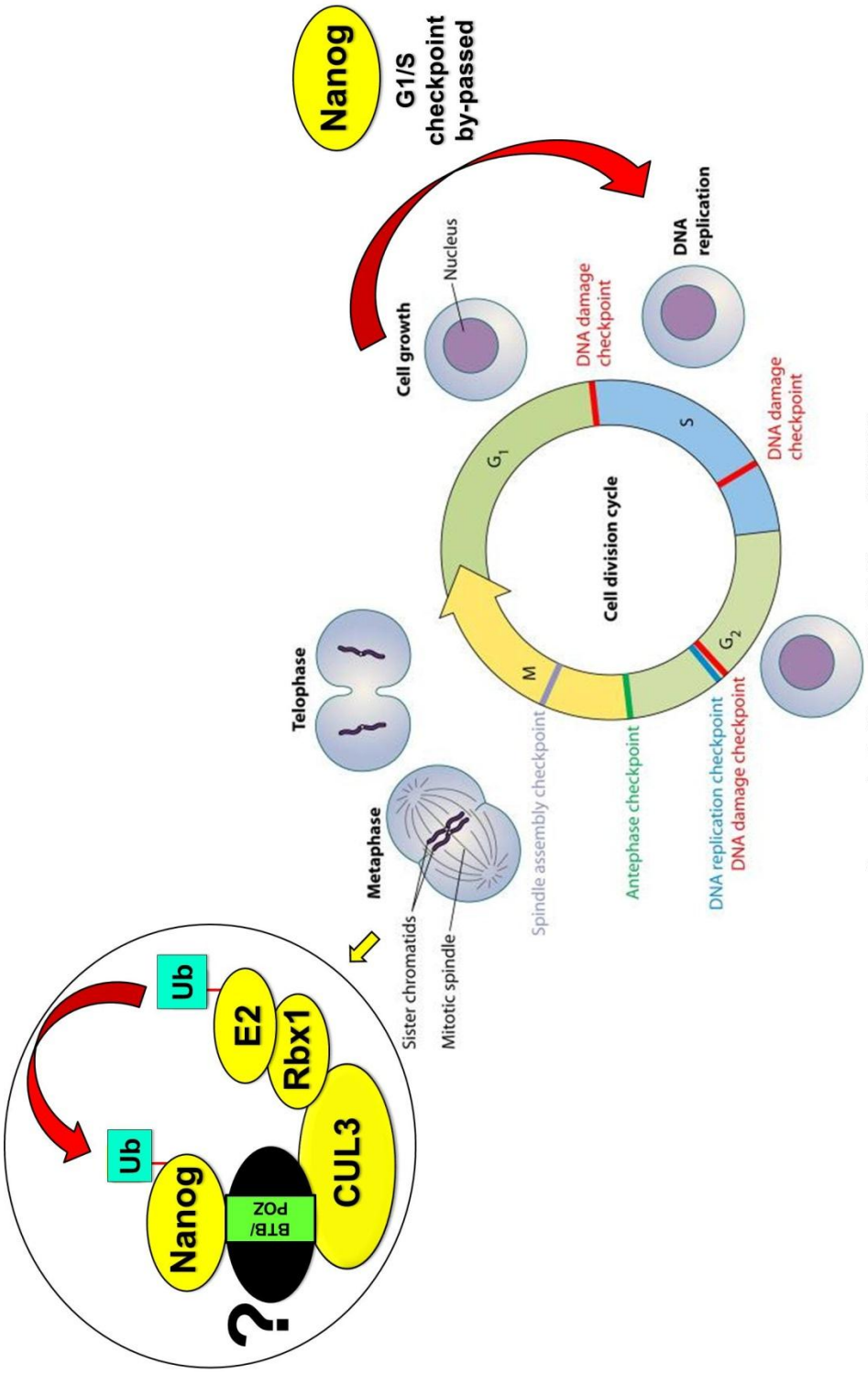


Image adapted from *Chin CF, Yeong FM* (2010).

Figure 4.20 Model of Cell cycle-dependent CUL3 CRL-mediated degradation of Nanog. Nanog is targeted by a CUL3 CRL (identified in this study) for ubiquitination and subsequent degradation during mitosis. Question mark indicates that the BTB/POZ domain-containing protein linking CUL3 to Nanog is currently unknown. Nanog protein accumulates during G1 phase and elevated levels of Nanog expressed at G1 phase, facilitates the G1 to S phase transition, cells by-pass the G1/S checkpoint (Zhang *et al.*, 2009; This study).Highlighted in yellow are those components of the model identified in this study.

4.4.11 Model of Cell Cycle-dependent CUL4A/B CRL-mediated degradation of Sox2 during mitosis:

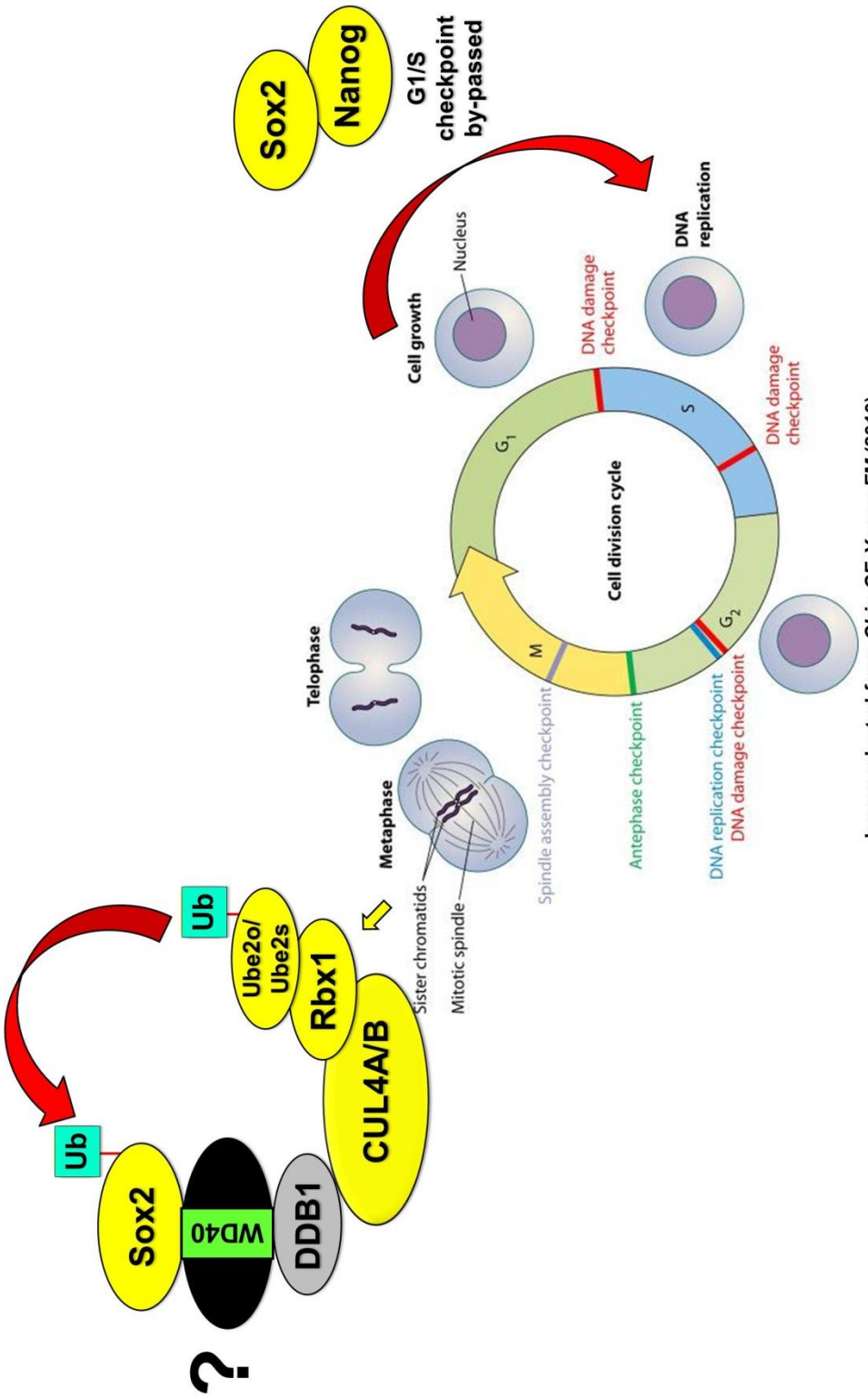


Image adapted from *Chin CF, Yeong FM* (2010).

Figure 4.21 Model of Cell cycle-dependent CUL4A/B CRL-mediated degradation of Sox2 during mitosis. Sox2 is targeted by a CUL4A/B CRL (identified in this study) for ubiquitination and subsequent degradation during mitosis. Question mark indicates that the WD40 domain-containing protein linking either CUL4A or CUL4B to Sox2 is currently unknown. Sox2 protein re-accumulates during G1 phase along with Nanog. Elevated levels of Sox2, in turn elevates Nanog protein levels with whom it interacts to help facilitate the G1 to S phase transition, cells by-pass the G1/S checkpoint (Zhang *et al.*, 2009; This study). Highlighted in yellow are those components of the model identified in this study.

4.4.11 Model of CUL4B isoform-specific CRL-mediated degradation of cytoplasmic Sox2:

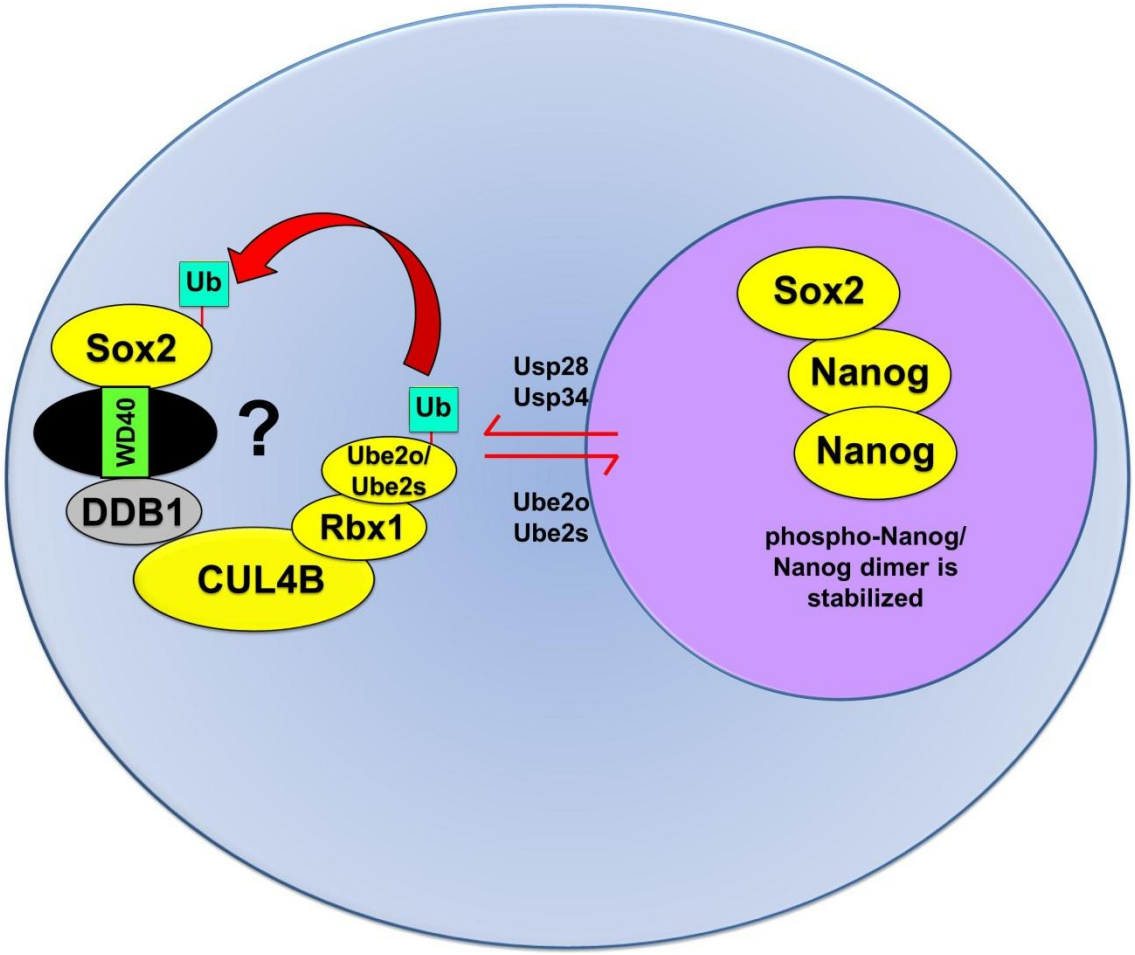


Figure 4.22 Model of CUL4B isoform-specific CRL-mediated Ubiquitination/ degradation of cytoplasmic Sox2. Cytoplasmic Sox2 is reversibly targeted by a CUL4B isoform-specific CRL (identified in this study) for ubiquitination and subsequent degradation. Question mark indicates that the WD40 domain-containing protein linking CUL4B to Sox2 is currently unknown. This reaction is reversed by the ubiquitin peptidases responsible for de-ubiquitinating Sox2 resulting in its being shuttled back into the nucleus. Highlighted in yellow are those components of the model identified in this study.

4.5 Future work:

4.5.1 Molecules identified in this study to close the molecular gateway toward differentiation.

Two molecules were identified in this study to increase phospho-Nanog levels and increase its interaction with Sox2 in embryonic stem cells are: Ht-31 and PKI (14-22). As described in section 4.4.3, they are both commercially available cell permeable peptides that inhibit PKA signaling by two distinct and highly specific mechanisms. In this study, it was shown that these two peptides have the ability to increase the levels of phospho-Nanog which results in its self-dimerization (Mullin *et al.*, 2008). This self-dimerization is necessary to maintain pluripotency and increase Nanog's transcriptional activity (Wang *et al.*, 2008). This is the first report of molecules capable of increasing phospho-Nanog levels and stabilizing it.

These peptides should be tested for their ability to maintain ES cells in a pluripotent state for longer time periods in culture both with and without LIF present in the culture conditions. Short-term culture in the presence of these molecules revealed that they out-perform the well-known 3*i* medium in their ability to increase phospho-Nanog levels. Therefore, these molecules should be tested for their ability to maintain ES cells in long-term culture and in direct comparison with 3*i* medium (Ying *et al.*, 2008). 3*i* medium requires the complete removal of FBS from medium, which is not required in this case. While the 3*i* system is currently the gold standard for maintaining pluripotency in ES cells, it has its drawbacks. For example, switching from standard culture conditions to serum-free 3*i* medium means that all previously obtained results may not be comparable or reproducible. Due to its cost, 3*i* medium's widespread use is restricted to low-scale culture of ES cells in academic research labs.

The advantage of using these cell permeable peptides is that, unlike 3*i* medium, they work in standard culture conditions (Figure 4.1) and can simply be added to standard culture medium currently being used by any lab.

4.5.2 Molecules identified in this study to open the molecular gateway toward differentiation.

In this study the chemical compound H89, which is widely marketed as PKA inhibitor, was shown to deplete the levels of phospho-Nanog and cause Nanog complex dissociation and degradation. When compared with a highly specific form of PKA inhibition using the cell permeable peptide PKI (14-22), however, the exact opposite effect was observed. That is, specific inhibition of PKA increased phospho-Nanog levels.

It has been well documented that H89 inhibits up to 8 other kinases in addition to PKA. They are: MAPKAP-K1b, MSK1, PKBα/Akt1, SGK, S6K1, ROCK II, AMPK, and CHK1 (Davies *et al.*, 2000; Lochner *et al.*, 2006). MSK1, S6K1 and ROCKII are inhibited with an IC50 of 120 nM, 80 nM, 270 nM, respectively. In the cases of MSK1 and S6K1, they are actually inhibited more strongly than PKA itself which has an IC50 of 135 nM. Inhibitors to all 8 kinases, and combinations thereof, should be tested for their ability to deplete phospho-Nanog levels and induce differentiation. The identification of such a compound would not only provide a powerful tool for regenerative medicine but also a potent anti-cancer therapeutic.

4.5.3 Investigating the Sox2 interaction with dimerized Nanog.

In this chapter, the interaction between Sox2 and the dimerized form of Nanog (phospho-Nanog) is demonstrated in mESCs. Also shown in this study is that this interaction is enriched for when Nanog is stabilized by an increase in phospho-Nanog levels. This implies that Sox2 interacts with the phosphorylated and dimerized forms Nanog with a higher affinity than its unmodified counterparts. This, in turn, results in higher nuclear Sox2 levels (shown in this study).

Taken together, these results suggest that Sox2 is part of the protein complex along with Nanog that maintains pluripotency. However, the functional significance of this interaction should be further studied. For example, Sox2 shRNA should be employed and the levels of phospho-Nanog determined. A decrease in phospho-Nanog levels would imply that the interaction with Sox2 somehow stabilizes the Nanog dimer. Conversely, an increase would imply that Sox2 somehow attenuates Nanog dimerization.

4.5.4 Identification of the BTB/POZ domain containing protein that interacts with Nanog and targets it for ubiquitination.

CUL3 was shown for the first time in this study as being involved for the degradation of Nanog. CUL3 binds directly to its specific substrate proteins via a BTB/POZ domain containing protein. In order to identify the BTB/POZ domain-containing protein that forms the CUL3 CRL degrading Nanog, a search for all BTB/POZ domain-containing interactors identified in the Nanog interaction network in this study should be performed. Each of the identified BTB/POZ domain containing proteins should then be tested for its ability to co-purify with Nanog. Subsequently, an *in vitro* ubiquitination assay using purified CUL3, Nanog and the

yet to be determined BTB/POZ domain containing protein should be performed. Additionally, the mitotic shake-off experiment should be repeated in order to accurately determine which stage of the cell cycle Nanog is being degraded by CUL3 in conjunction with the unknown BTB-POZ domain-containing protein.

4.5.5 Identification of the CUL4B substrate-specific adaptor protein that targets Sox2 for ubiquitination.

For the first time in this study, the CUL4B isoform of CUL4 was identified as being involved in the degradation of Sox2. The best known CUL4 E3 ligases are multi-subunit complexes containing CUL4, ROC1/RBX1 and DDB1. DDB1 binds to a WD40 repeat containing protein which serves as the substrate-specific adaptor protein. These ligases are involved in DNA damage repair and cell cycle progression. Although ROC1/RBX1 was identified by mass spectrometry in this study as an interactor with Sox2, DDB1 was not identified significantly above the control IP following the high stringency selection criteria used in this study. This could be due to the fact that the criteria used in this study to define interactors may have been too stringent.

In addition, since CUL4B was observed to be primarily cytoplasmic in ES cells and translocated to the nucleus only upon entry into mitosis, where CUL4A is capable of substituting for CUL4B function. This observation leads to the following hypothesis: That Sox2 is degraded in the nucleus by a CUL4A/B –ROC1-DDB1-WD40 domain-containing protein X during mitosis. In order to identify protein X, the Sox2 interaction data set from this study should be searched for all WD40 repeat containing proteins. Each protein identified would then be tested for their ability to co-purify with Sox2. Once identified, an *in vitro* ubiquitination assay using purified

Sox2, ROC1/RBX1, DDB1 and the yet to be identified WD40 repeat-containing protein should be performed. If this is unsuccessful an IP-LC-MS/MS should be performed on HA-Sox2 cells synchronized in mitosis in an attempt to enrich for this low abundant ligase.

CHAPTER 5: Thesis Summary/Conclusion

5.1 The IP-LC-MS/MS Protein-Protein Interaction Networks in mESCs

The data presented in this thesis originated from a desire to identify post-translational mechanisms that govern pluripotency. This was achieved by employing an IP-LC-MS/MS strategy developed specifically for the identification of proteins that interact with Nanog, Sox2, Esrrβ and Sall4. By employing this strategy, novel protein interactions unique to this study were observed. This has provided a very different view of the protein interaction landscape in mESCs compared to that from the 10 previously reported studies summarized by Mallana and Rizzino, 2012.

While high salt buffers are commonly used when extracting transcription factors and nuclear proteins from cells, these high ionic strength conditions also disrupt weak protein-protein interactions. Since the goal of this study was to identify the weak and typically transient interactions that occur between a kinase or phosphatase and its target, lower ionic strength buffers (250 mM NaCl) were employed at all steps during immunopurifications. Although lower salt containing buffers were used, Benzonase as well as the physical shearing of cells with pipette tips were employed to ensure adequate lysis of both cells and nuclei.

The interaction networks described in this thesis have extended our view of the protein-protein interactions in which these pluripotency factors engage with a variety of kinases, phosphatases and E3 ubiquitin ligases, indicating enzymes that may be modulate the abundance or activity of pluripotency factors. As a result, the protein interaction networks identified in this study have perhaps laid the foundation

for future work to identify means by which the stabilization or destabilization of pluripotency may be achieved.

5.2 Nanog Interacting Proteins

The protein interaction network ascertained for Nanog under the lower ionic strength conditions used in this study contains several novel interactors including Sox2, Esrrβ, Pik3cd, BRCA2 and CUL4B. Additionally, several proteins responsible for performing various types of post-translational modifications were identified to interact with Nanog. In total, 7 kinases, 2 cullins, 5 E3 ubiquitin ligases and 1 F-box protein were identified. The Nanog protein interaction network presented in this thesis is the first to identify these protein-modifying enzymes as potential Nanog interactors/modulators.

5.2.1 BRCA2

BRCA2, a known tumour suppressor protein whose best characterized function is its involvement in error-free repair of double-stranded DNA breaks through homologous recombination and the (Duncan *et al.*, 1998; Sarkisian *et al.*, 2001; Yoshida and Miki, 2004), was also found to be an interactor. This is an entirely novel aspect of the present study and suggests a role for Nanog in the maintenance of the genomic integrity of ES cells. Furthermore, this hints that the gateway to pluripotency may be mediated, in part, by DNA damage and repair.

It has been previously reported that both DNA damage and differentiation significantly decrease when cells are grown under hypoxic conditions (Di Carlo *et al.*, 2004; Moore and Lemishka, 2006; Pires *et al.*, 2010) due to a decrease in the

production of Reactive Oxygen Species (ROS) (Evans and Cooke, 2004; Bertram and Hass, 2008; Guachalla and Rudolph, 2010) . In addition, a recent body of literature (Larsen *et al.*, 2010a; Larsen *et al.*, 2010b; Milyavsky *et al.*, 2010) points to the requirement for adult stem cells to damage their own DNA in order to reveal genetic elements deeply buried within the structure of DNA and allow transcriptional programs associated with differentiation to become activated (Fernando and Megeney, 2007; Abdul-Ghani and Megeney, 2008; Larsen et al., 2010b; Milyavsky et al., 2010). Could this also be true for embryonic stem cells?

The observed Nanog-BRCA2 interaction in this study lays the foundation for a rather appealing hypothesis to be tested in future work.

5.2.2 The Nanog-BRCA2 Hypothesis

BRCA2 was identified as an interactor with Nanog in normoxic conditions (Chapter 3 of this study) when levels of DNA damage-inducing ROS are high. Therefore, it was hypothesized that this interaction would be less pronounced in hypoxia when ROS-induced DNA damage levels are lowered. This hypothesis is illustrated in Figure 5.1. The red arrows indicate that the Nanog-BRCA2 interaction is either promoted (required) or inhibited (no longer required). The red question mark indicates that it is not known whether or not the Nanog-BRCA2 interaction is reduced by lowered oxygen tension.

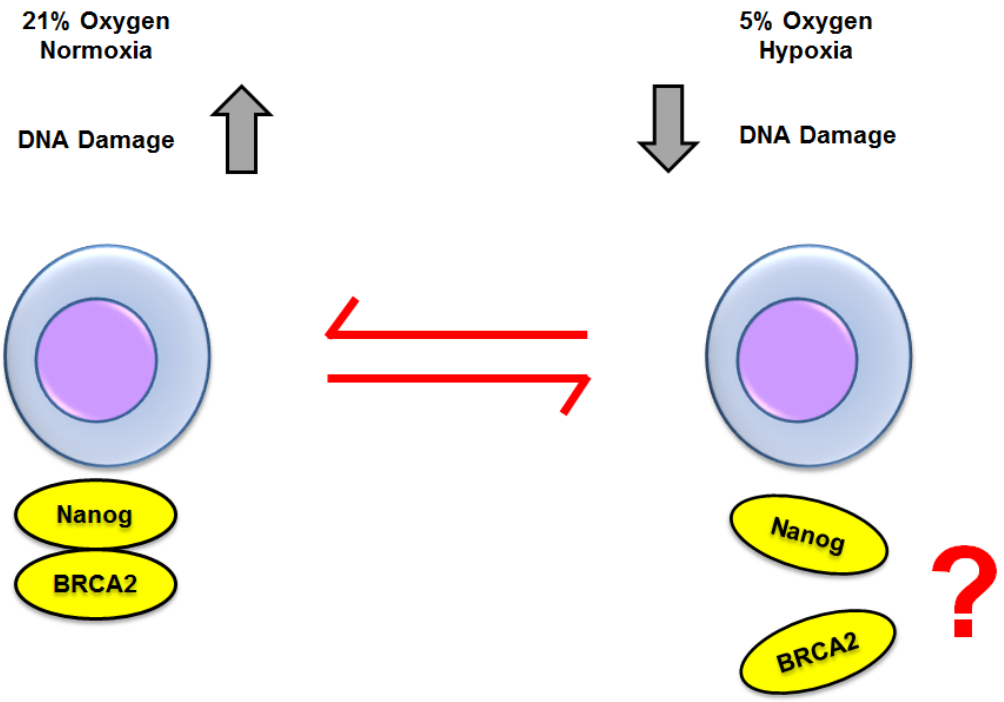


Figure 5.1: Model of Nanog interaction with BRCA2 in order to facilitate DNA damage repair and prevent differentiation of ES cells. In this model, Nanog and BRCA2 participate in DNA damage repair due to oxidative stress. In physiological conditions, this interaction is no longer needed as DNA damage ceases due to lowered oxygen tension.

5.2.3 The BRCA2 and Hmg20b complex

In addition to its role in DNA damage repair, BRCA2 is also known to regulate entry into mitosis through its interaction with Hmg20b. Hmg20b was first identified as BRAF35 (BRCA2 Associated Factor 35), the DNA-binding component of a complex it forms with the BRCA2 protein. Together, Hmg20b and BRCA2 were shown to be involved in cell cycle progression at the G2/M transition (Marmorstein *et al.*, 2001). Hmg20b is also part of a complex devoid of BRCA2 in which it associates with several histone deacetylase proteins and participates in the repression of neuronal-specific genes through chromatin remodeling (Hakimi *et al.*, 2002;

Hakimi *et al.*, 2003; Iwase *et al.*, 2004). In Chapter 3, both BRCA2 and its interacting partner Hmg20b (in the Esrrβ protein interaction network) were identified as part of the integrated interaction network in this work. The identification of both BRCA2 and Hmg20b may indicate a role for the Nanog-BRCA2 interaction in cell cycle progression, chromatin remodeling, neuronal gene-specific transcriptional repression and DNA repair. Further investigation of this complex and its association with Nanog may uncover unknown functions of these proteins in the maintenance of pluripotency.

5.3 Modulation of Nanog and Sox2 Using Small Molecules and Cell Permeable Peptides

In Chapter 4 of this work, a chemical biology approach was taken in order to identify interrogate the potential action of the identified kinases and Cullin-RING ligases (CRLs) on Nanog. Novel to this work is the identification of several small molecules capable of either elevating phospho-Nanog levels (MLN4924, PKI (14-22) amide, Aurk*i*III, Ht-31 and Nocodazole) or depleting phospho-Nanog levels (H89, Wortmannin and Forskolin) in mESCs.

Since phosphorylation of Nanog has been reported to stabilize Nanog and induce Nanog-Nanog dimer formation (Moretto-Zita *et al.*, 2010), it is postulated here that some of the compounds identified in this study may stabilize phospho-Nanog levels, induce dimer formation and close the gateway to differentiation. This will require to be demonstrated by further experimentation in which mESCs treated with the phospho-Nanog elevating compounds identified here are induced to differentiate. A delay in the induction of mESC differentiation would functionally validate the ‘Molecular Keys to the Gateway of Pluripotency’ model proposed in this

thesis.

Although the IP-LC-MS/MS studies performed by both Pardo (Pardo *et al.*, 2010) and Van den Berg (Van den Berg *et al.*, 2010) identified Akap8 (Protein kinase A anchoring protein) as part of their protein interaction networks, neither study identified the kinase itself. Novel to this study was the identification of several components of the PKA signaling pathway as part of the integrated protein interaction network. Among them were, Akap8 and the regulatory subunits of PKA itself. Using small molecules and sequence-specific cell permeable peptides, it was shown in this study that the specific inhibition of components of the PKA signaling pathway modulates both Nanog and Sox2 proteins in mESCs.

While preparing to defend this thesis, a work by Yamimizu *et al.*, 2012 was published in Cell Stem Cell. In this work, Yamamizu and colleagues demonstrate that specifically inhibiting PKA using PKI (14-22) amide delays differentiation of mESCs by maintaining elevated Nanog transcript levels. However, the ability of PKI (14-22) to elevate phospho-Nanog protein levels in mESCs currently remains novel to this thesis.

5.4 Modulation of Nanog and Sox2 by Inhibiting Cullin RING Ligase Activity in mESCs

Among the compounds identified to elevate phospho-Nanog levels in this work was the pan-Cullin RING Ligase inhibitor, MLN4924. This raises the possibility that there may be several phosphorylated forms of Nanog that exist in mESCs. That is, phospho-Nanog proteins which are stabilized and may form Nanog-Nanog dimers and other forms which are ubiquitinated by an unidentified CRL and subsequently degraded by the 26S proteasome.

In an effort to identify the exact CRL involved in the degradation of Nanog, each cullin protein identified as being part of the integrated protein interaction network in this study was inhibited using shRNA. Specific shRNA inhibition of CUL3 resulted in a 100% increase in Nanog levels as well as causing Nanog protein to be expressed 8 hours earlier than in control shRNA cells following release from mitotic arrest. These results are the first demonstration that Nanog may be targeted by a CRL E3 ubiquitin ligase in a cell cycle-dependent manner.

Also in this work, it is shown that Sox2 protein levels are increased by 190% in response to CUL4B inhibition by shRNA. Taken together with the observed timing of Nanog expression in the CUL4B shRNA cells, it would appear that Sox2 is also being degraded in a cell cycle-dependent manner similar to Nanog. This is the first demonstration that Sox2 may be degraded by a CRL E3 Ubiquitin ligase.

5.5 Co-localization of Nanog, Sox2 and CUL4A/B in mESCs

In Chapter 4 it was shown that Nanog and Sox2 proteins co-localize with each other and that this co-localization appears to be strongest during mitosis on mitotic spindles and condensed chromosomes. Furthermore, it is shown in this work that Sox2 and CUL4A/B immunostaining occurs primarily in the cytoplasm with the exception of mitosis, where both protein localize to the nucleus.

While a cytoplasmic localization for Sox2 has been reported by others, (eg., Baltus *et al.*, 2009), it remains possible that this immunostaining is non-specific. For example, it is possible that the blocking steps performed during immunostaining were unsuccessful and incomplete resulting in high levels of background staining in the cytoplasm. It is also possible that the primary or secondary antibodies themselves

recognize other epitopes than those they were intended to bind. This explanation seems highly unlikely since the immunostaining experiments on the 3XFLAG-Nanog cell line as well as the CUL4B shRNA cell line were performed using exactly the same protocol, primary and secondary antibodies.

One way to address concerns regarding the specificity of the antibodies used would be to perform the same immunostaining experiments on the CUL4B shRNA cell line and the parental control line using a Sox2 antibody both with and without a Sox2 epitope blocking peptide. The disappearance of Sox2 signal from the cytoplasm and/or mitotic spindles will indicate that the immunofluorescence signal in the cytoplasm and at mitotic spindles is indeed indicative of Sox2. Another control for non-specific interaction of the Sox2 antibody used would be to perform Sox2 immunostaining on a Sox2 shRNA cell line. If a large amount of immunofluorescence is detected in the cytoplasm in this cell line, it would clearly indicate a lack of specificity by the antibody employed in this study. In order to determine the non-specific interactions by the secondary antibodies used in this study, cells should be incubated with the secondary antibodies alone. This type of experiment will reveal whether or not the fluorescent signals observed in this work are the result of the secondary antibodies binding to components of the blocking reagent employed.

Yet another experimental control would be to perform anti-HA immunostaining on the 6XHA-Sox2 cell line developed in this study. Immunostaining could be performed both in the presence and absence of the 6XHA peptide. This experiment would reveal the cellular localization of the 6XHA-Sox2

protein used to ascertain the Sox2 protein interaction network by IP-LC-MS/MS in this study.

With regards to the cytoplasmic immunostaining observed for CUL4A/B in this study, further experiments should be performed using the CUL4B-specific antibody developed by Guerrero-Santoro *et al.*, 2006. Additionally, by immunostaining for CUL4B in the presence of the CUL4B blocking peptide that Guerrero-Santoro *et al.*, 2006 used, the disappearance of CUL4B signal will indicate where the CUL4B protein is localized in mESCs.

While some future work remains to be performed with regards to immunofluorescence controls, the identification here that the specific inhibition of the CUL4B isoform elevates Sox2 levels provides a novel insight into the post-translational regulation of this pluripotency factor. Furthermore, the data presented in Chapter 4 of this thesis provides evidence that the abundance of Nanog and Sox2 proteins may be modulated by non-genetic means in mESCs. The novel pathways and chemical inhibitors identified in this work have further extended of knowledge of the signaling cascades that govern the maintenance of pluripotency in ES cells.

REFERENCES

Abdul-Ghani M, Megeney LA (2008). Rehabilitation of a contract killer: caspase-3 directs stem cell differentiation. *Cell Stem Cell* **2**(6):515-6.

Akasaka T, Takahashi N, Suzuki M, Koseki H, Bodmer R, Koga H (2002). MBLR, a new RING finger protein resembling mammalian Polycomb gene products, is regulated by cell cycle-dependent phosphorylation. *Genes Cells* **7**(8):835-50.

Amerik AY, Hochstrasser M (2004). Mechanism and function of deubiquitinating enzymes. *Biochim Biophys Acta* **1695**(1-3):189-207.

Andrews P, He YJ, Xiong Y (2006). Cytoplasmic localized ubiquitin ligase cullin 7 binds to p53 and promotes cell growth by antagonizing p53 function. *Oncogene* **25**(33):4534-48.

Angers S, Li T, Yi X, MacCoss MJ, Moon RT, Zheng N (2006). Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* **443**(7111):590-3.

Ardley HC, Robinson PA. E3 ubiquitin ligases (2005). *Essays Biochem* **41**:15–30.

Babaie Y, Herwig R, Greber B, Brink TC, Wruck W, Groth D, Lehrach H, Burdon T, Adjaye J. (2007). Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. *Stem Cells* **25**(2):500-10.

Ballabeni A, Park IH, Zhao R, Wang W, Lerou PH, Daley GQ, Kirschner MW Cell cycle adaptations of embryonic stem cells (2011). *Proc Natl Acad Sci U S A* **108**(48):19252-7.

Baltus GA, Kowalski MP, Zhai H, Tutter AV, Quinn D, Wall D, Kadam S (2009). Acetylation of sox2 induces its nuclear export in embryonic stem cells. *Stem Cells* **27**(9):2175-84.

Beddington RSP and Robertson EJ (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development 105, 733–737.

Bell CE, Larivière NM, Watson PH, Watson AJ (2009). Mitogen-activated protein kinase (MAPK) pathways mediate embryonic responses to culture medium osmolarity by regulating Aquaporin 3 and 9 expression and localization, as well as embryonic apoptosis. *Hum Reprod* **24**(6):1373-86.

Bertram C, Hass R (2008). Cellular responses to reactive oxygen species-induced DNA damage and aging. *Biol Chem* **389**(3):211-20.

Bolos V, Grego-Bessa J, and de la Pompa JL (2007) Notch signalling in Development and cancer. *Endocr Rev* **28**, 339–363.

Borowiak M, Maehr R, Chen S, Chen AE, Tang W, Fox JL, Schreiber SL, Melton DA (2009). Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. *Cell Stem Cell* **4**(4):348-58.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**:947–956.

Brandner S (2010). Nanog, Gli, and p53: a new network of stemness in development and cancer. *EMBO J* **29**(15):2475-6.

Burdon T, Smith A, Savatier P (2002). Signalling, cell cycle and pluripotency in embryonic stem cells. *Trends Cell Biol* **12**(9):432-8.

Burdon T, Chambers I, Stracey C, Niwa H, Smith A (1999). Signaling mechanisms regulating self-renewal and differentiation of pluripotent embryonic stem cells. *Cells Tissues Organs* **165**(3-4):131-43.

Burrows JF, Johnston JA (2012). Regulation of cellular responses by deubiquitinating enzymes: an update. *Front Biosci* **17**:1184-200.

Carr DW, Stofko-Hahn RE, Fraser ID, Bishop SM, Acott TS, Brennan RG, Scott JD (1991). Interaction of the regulatory subunit (RII) of cAMP-dependent protein kinase with RII-anchoring proteins occurs through an amphipathic helix binding motif. *J Biol Chem* **266**(22):14188-92.

Chae, H.-D., Lee, M.-R. and Broxmeyer, H. E. (2012). 5-Aminoimidazole-4-carboxyamide Ribonucleoside Induces G_1/S Arrest and Nanog Downregulation via p53 and Enhances Erythroid Differentiation. *Stem Cells* **30**: 140–149.

Chambers I, Colby D, Robertson M, et al. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* 113:643–655.

Chambers I, Smith A (2004). Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene* **23**(43):7150-60.

Chambers I (2004). The molecular basis of pluripotency in mouse embryonic stem cells. *Cloning Stem Cells* **6**(4):386-91.

Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith AG (2007). Nanog safeguards pluripotency and mediates germline development. *Nature* **50**(7173):1230-4.

Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**:544–547.

Chen G, Gharib TG, Huang CC, Taylor JM, Misek DE, Kardia SL, Giordano TJ, Iannettoni MD, Orringer MB, Hanash SM, Beer DG (2002). Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics* **1**(4):304-13.

Chen GI, Gingras AC (2007). Affinity-purification mass spectrometry (AP-MS) of serine/threonine phosphatases. *Methods* **42**(3):298-305.

Chen S, Do JT, Zhang Q, Yao S, Yan F, Peters EC, Schöler HR, Schultz PG, Ding S (2006).Self-renewal of embryonic stem cells by a small molecule. *Proc Natl Acad Sci U S A* **103**(46):17266-71.

Chen S, Borowiak M, Fox JL, Maehr R, Osafune K, Davidow L, Lam K, Peng LF, Schreiber SL, Rubin LL, Melton D (2009). A small molecule that directs differentiation of human ESCs into the pancreatic lineage. *Nat Chem Biol* **5**(4):258-65.

Chen T, Du J, Lu G (2012). Cell growth arrest and apoptosis induced by Oct4 or Nanog knockdown in mouse embryonic stem cells: a possible role of Trp53. *Mol Biol Rep* **39**(2):1855-61.

Cheong CY, Lon Ng PM, Ponnampalam R, Tsai HH, Bourque G, Lufkin T (2011). In silico tandem affinity purification refines an Oct4 interaction list. *Stem Cell Res Ther* **2**(3):26.

Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, Li P, Ang YS, Lim B, Robson P, Ng HH (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* **25**(14):6031-46.

Chin CF, Yeong FM (2010). Safeguarding entry into mitosis: the antephase checkpoint. *Mol Cell Biol* **30**(1):22-32.

Chiou SH, Yu CC, Huang CY, Lin SC, Liu CJ, Tsai TH, Chou SH, Chien CS, Ku HH, Lo JF (2008). Positive correlations of Oct-4 and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma. *Clin Cancer Res***14**(13):4085-95.

Chiou SH, Wang ML, Chou YT, Chen CJ, Hong CF, Hsieh WJ, Chang HT, Chen YS, Lin TW, Hsu HS, Wu CW (2010). Coexpression of Oct4 and Nanog enhances malignancy in lung adenocarcinoma by inducing cancer stem cell-like properties and epithelial-mesenchymal transdifferentiation. *Cancer Res* **70**(24):10433-44.

Cox J, Mann M (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **12**:1367-72.

Cox JL, Rizzino A (2010).Induced pluripotent stem cells: what lies beyond the paradigm shift. *Exp Biol Med (Maywood)* **235**(2):148-58.

Craig KL, Tyers M (1999). The F-box: a new motif for ubiquitin dependent proteolysis in cell cycle regulation and signal transduction. *Prog Biophys Mol Biol* **72**(3):299-328.

Davies SP, Reddy H, Caivano M, Cohen P (2000). Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem J* **351**:95–105.

Deshaies R J and Joazeiro C A (2009). RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **78**, 399-434.

Di Carlo A, De Mori R, Martelli F, Pompilio G, Capogrossi MC, Germani A (2004).Hypoxia inhibits myogenic differentiation through accelerated MyoD degradation. *J Biol Chem* **16**;279(16):16332-8.

Ding S, Wu TY, Brinker A, Peters EC, Hur W, Gray NS, Schultz PG (2003). Synthetic small molecules that control stem cell fate. *Proc Natl Acad Sci U S A* **100**(13):7632-7.

Ding S, Schultz PG (2004). A role for chemistry in stem cell biology. *Nat Biotechnol* **22**(7):833-40.

Do JT, Han DW, Gentile L, Sobek-Klocke I, Stehling M, Lee HT, Schöler HR (2007). Erasure of cellular memory by fusion with pluripotent cells. *Stem Cells* **25**(4):1013-20.

Dolganov GM, Maser RS, Novikov A, Tosto L, Chong S, Bressan DA, Petrini JH (1996). Human Rad50 is physically associated with human Mre11: identification of a conserved multiprotein complex implicated in recombinational DNA repair. *Mol Cell Biol* **16**(9):4832-41.

Duncan JA, Reeves JR, Cooke TG (1998). BRCA1 and BRCA2 proteins: roles in health and disease. *Molecular pathology* **51**(5): 237–47.

Eminli S, Utikal J, Arnold K, Jaenisch R, Hochedlinger K (2008). Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells* **26**(10):2467-74.

Etlinger JD, Goldberg AL (1977). A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. *Proc Natl Acad Sci U S A* **74**(1):54-8.

Evans MD, Cooke MS (May 2004). Factors contributing to the outcome of oxidative damage to nucleic acids. *Bioessays* **26** (5): 533–42.

Evans MJ and Kaufman MH (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**:154-156.

Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H, Bergman Y (2006). G9amediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* **8**:188–194.

Feng B, Jiang J, Kraus P, Ng JH, Heng JC, Chan YS, Yaw LP, Zhang W, Loh YH, Han J, Vega VB, Cacheux-Rataboul V, Lim B, Lufkin T, Ng HH (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat Cell Biol* **11**(2):197-203.

Fernando P, Megeney LA (2007). Is caspase-dependent apoptosis only cell differentiation taken to the extreme? *FASEB J* **21**(1):8-17.

Furukawa M, Xiong Y (2005). BTB protein Keap1 targets antioxidant transcription factor Nrf2 for ubiquitination by the Cullin 3-Roc1 ligase. *Mol Cell Biol.* **25**(1):162-71.

Futcher B, Latter GI, Monardo P, McLaughlin CS, Garrels JI (1999). A sampling of the yeast proteome. *Mol Cell Biol* **19**(11):7357-68.

Gan L, Zheng W, Chabot JG, Unterman TG, Quirion R (2005). Nuclear/cytoplasmic shuttling of the transcription factor FoxO1 is regulated by neurotrophic factors J *Neurochem* **93**(5):1209-19.

Ganoth D, Leshinsky E, Eytan E, and Hershko A (1988). A multicomponent system that degrades proteins conjugated to ubiquitin. Resolution of components and evidence for ATP-dependent complex formation. *J Biol Chem* **263**, 12412-1241.

Gardner RL and Brook FA (1997). Reflections on the biology of embryonic stem cells. Int J Dev Biol 41, 235–243.

Geyer R, Wee S, Anderson S, Yates J, Wolf DA (2003). BTB/POZ domain proteins are putative substrate adaptors for cullin 3 ubiquitin ligases. *Mol Cell* **12**(3):783-90.

Glass DB, Cheng HC, Mende-Mueller L, Reed J, Walsh DA (1989).Primary structural determinants essential for potent inhibition of cAMP-dependent protein kinase by inhibitory peptides corresponding to the active portion of the heat-stable inhibitor protein. *J Biol Chem* **264**(15):8802-10.

Glickman MH, Ciechanover A (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev* **82**(2):373-428. Review.

Gong L, Yeh ET (1999). Identification of the activating and conjugating enzymes of the NEDD8 conjugation pathway. *J Biol Chem* **274**: 12036–12042.

Graf T, Stadtfeld M (2008). Heterogeneity of embryonic and adult stem cells. Cell *Stem Cell* **3**:480–483.

Gray WM, Hellmann H, Dharmasiri S, Estelle M (2002). Role of the Arabidopsis RING-H2 protein RBX1 in RUB modification and SCF function. *Plant Cell* **14**: 2137–2144.

Green EM, Gozani O. (2011). CUL4B: trash talking at chromatin. *Mol Cell* **43**(3):321-3.

Griffin TJ, Gygi SP, Ideker T, Rist B, Eng J, Hood L, Aebersold R (2002). Complementary profiling of gene expression at the transcriptome and proteome levels in Saccharomyces cerevisiae. *Mol Cell Proteomics* **1**(4):323-33.

Goldknopf IL, Busch H (1977). Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. *Proc Natl Acad Sci U S A* **74**(3):864-8.

Groisman R, Polanowska J, Kuraoka I, Sawada J, Saijo M, Drapkin R, Kisselev AF, Tanaka K, Nakatani Y (2003) The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell* **113**:357–367.

Guachalla LM, Rudolph KL (2010). ROS induced DNA damage and checkpoint responses: influences on aging? *Cell Cycle* **9**(20):4058-60.

Guerrero-Santoro J, Kapetanaki MG, Hsieh CL, Gorbachinsky I, Levine AS, Rapić-Otrin V (2008). The cullin 4B-based UV-damaged DNA-binding protein ligase binds to UV-damaged chromatin and ubiquitinates histone H2A. *Cancer Res* **68**(13):5014-22.

Guo Y, Nady N, Qi C, Allali-Hassani A, Zhu H, Pan P, Adams-Cioaba MA, Amaya MF, Dong A, Vedadi M, Schapira M, Read RJ, Arrowsmith CH, Min J (2009). Methylation-state-specific recognition of histones by the MBT repeat protein L3MBTL2. *Nucleic Acids Res* **37**(7):2204-10.

Gurdon JB, Bourillot PY (2001). Morphogen gradient interpretation. *Nature* **413**:797–803.

Gygi SP, Rochon Y, Franza BR, Aebersold R (1999). Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* **19**(3):1720-30.

Hakenjos JP, Richter R, Dohmann EM, Katsiarimpa A, Isono E, Schwechheimer C (2011). MLN4924 is an efficient inhibitor of NEDD8 conjugation in plants. *Plant Physiol* **56**(2):527-36.

Hakimi MA, Bochar DA, Chenoweth J, Lane WS, Mandel G, Shiekhattar R.
A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes (2002). *Proc Natl Acad Sci U S A* **99**(11):7420-5.

Hakimi, Mohamed-Ali, Dong Yuanshu, Lane William S, Speicher David W, Shiekhattar Ramin (2003). A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. *J Biol Chem* **278**(9): 7234–9.

Haas AL, Warms JV, Hershko A, Rose IA (1982). Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *J Biol Chem* **257**(5):2543-8.

Hack CJ (2004). Integrated transcriptome and proteome data: the challenges ahead. *Brief Funct Genomic Proteomic* **3**(3):212-9.

Han MK, Song EK, Guo Y, Ou X, Mantel C, Broxmeyer HE (2008). SIRT1 regulates apoptosis and Nanog expression in mouse embryonic stem cells by controlling p53 subcellular localization. *Cell Stem Cell* **2**(3):241-51.

Hart AH, Hartley L, Parker K, Ibrahim M, Looijenga LH, Pauchnik M, Chow CW, Robb L (2005). The pluripotency homeobox gene NANOG is expressed in human germ cell tumors. *Cancer* **104** (10): 2092–8.

Hatakeyama S, Nakayama KI (2003). U-box proteins as a new family of ubiquitin ligases. *Biochem Biophys Res Commun* **302**:635–45.

Helbig AO, Daran-Lapujade P, van Maris AJ, de Hulster EA, de Ridder D, Pronk JT, Heck AJ, Slijper M (2011). The diversity of protein turnover and abundance under nitrogen-limited steady-state conditions in Saccharomyces cerevisiae. *Mol Biosyst* **7**(12):3316-26.

Heng JC, Feng B, Han J, Jiang J, Kraus P, Ng JH, Orlov YL, Huss M, Yang L, Lufkin T, Lim B, Ng HH (2010). The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell* **6**(2):167-74.

Hershko A, Ciechanover A, Heller H, Haas AL, and Rose IA (1980). Proposed role of ATP in protein breakdown: Conjugation of proteins with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc Natl Acad Sci USA* **77**, 1783–1786.

Hershko A, Heller H, Elias S and Ciechanover A (1983). Components of ubiquitin-protein ligase system: resolution, affinity purification and role in protein breakdown. *J Biol Chem* **258**, 8206–8214.
Hershko A, Ciechanover A (1998). The Ubiquitin System. *Annu. Rev. Biochem.* **67**:425–79.

Higa LA, Wu M, Ye T, Kobayashi R, Sun H, Zhang H (2006a). CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat Cell Biol* **8**(11):1277-83.

Higa LA, Banks D, Wu M, Kobayashi R, Sun H, Zhang H (2006b). L2DTL/CDT2 interacts with the CUL4/DDB1 complex and PCNA and regulates CDT1 proteolysis in response to DNA damage. *Cell Cycle* **5**(15):1675-80.

Hochstrasser M (2009). Origin and function of ubiquitin-like proteins. *Nature* **458**: 422–429.

Hoei-Hansen CE, Almstrup K, Nielsen JE, Brask Sonne S, Graem N, Skakkebaek NE, Leffers H, Rajpert-De Meyts E (2005). Stem cell pluripotency factor NANOG is expressed in human fetal gonocytes, testicular carcinoma in situ and germ cell tumours. *Histopathology* **47** (1): 48–56.

Hogan BL (1999). Morphogenesis. *Cell* **96**:225–233.

Hooper M, Hardy K, Handyside A, Hunter S, Monk M (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* **326**(6110):292-5.

Hori T, Osaka F, Chiba T, Miyamoto C, Okabayashi K, Shimbara N, Kato S, Tanaka K (1999). Covalent modification of all members of human cullin family proteins by NEDD8. *Oncogene* **18**: 6829–6834.

Hotton SK, Callis J (2008). Regulation of cullin RING ligases. *Annu Rev Plant Biol* **59**: 467–489.

Hrecka K, Gierszewska M, Srivastava S, Kozaczkiewicz L, Swanson SK, Florens L, Washburn MP, Skowronski J (2007). Lentiviral Vpr usurps Cul4-DDB1[VprBP] E3 ubiquitin ligase to modulate cell cycle. *Proc Natl Acad Sci* **104**(28):11778-83.

Huang W, Zhou X, Lefebvre V, de Crombrugghe B (2000). Phosphorylation of SOX9 by cyclic AMP-dependent protein kinase A enhances SOX9's ability to transactivate a Col2a1 chondrocyte-specific enhancer. *Mol Cell Biol* **20**, 4149-58.

Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S, Muhlestein W, Melton DA (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol* **26**(11):1269-75.

Hübner K, Fuhrmann G, Christenson LK, Kehler J, Reinbold R, De La Fuente R, Wood J, Strauss JF 3rd, Boiani M, Schöler HR (2003). Derivation of oocytes from mouse embryonic stem cells. Science 300(5623):1251-6.

Humphrey-Smith I, Cordwell S J and Blackstock W P (1997). Proteome research: Complementarity and limitations with respect to the RNA and DNA worlds. *Electrophoresis* **180**: 1217–1242.

Ichida JK, Blanchard J, Lam K, Son EY, Chung JE, Egli D, Loh KM, Carter AC, Di Giorgio FP, Koszka K, Huangfu D, Akutsu H, Liu DR, Rubin LL, Eggan K (2009). A small-molecule inhibitor of tgf-Beta signaling replaces sox2 in reprogramming by inducing nanog. Cell Stem Cell 6; 5(5):491-503.

Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L (2001). Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292**(5518):929-34.

Insel PA, Ostrom RS (2003). Forskolin as a tool for examining adenylyl cyclase expression, regulation, and G protein signaling. *Cell Mol Neurobiol* **23**(3):305-14.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. *Nature* **442**: 533-538.

Iwase, Shigeki, Januma Aya, Miyamoto Kiyoko, Shono Naomi, Honda Arata, Yanagisawa Junn, Baba Tadashi (2004). Characterization of BHC80 in BRAF-HDAC complex, involved in neuron-specific gene repression. *Biochem Biophys Res Commun* **322** (2): 601–8.

Iyer LM, Koonin EV, Aravind L (2004). Novel predicted peptidases with a potential role in the ubiquitin signaling pathway. *Cell Cycle* **3**(11):1440-50.

Jackson S, Xiong Y (2009). CRL4s: the CUL4-RING E3 ubiquitin ligases. *Trends Biochem Sci* **34**(11):562-70.

Jeong CH, Cho YY, Kim MO, Kim SH, Cho EJ, Lee SY, Jeon YJ, Lee KY, Yao K, Keum YS, Bode AM, Dong Z (2010). Phosphorylation of sox2 cooperates in reprogramming to pluripotent stem cells. Stem Cells 28, 2141-50.

Jeter CR, Badeaux M, Choy G, Chandra D, Patrawala L, Liu C, Calhoun-Davis T, Zaehres H, Daley GQ, Tang DG (2009). Functional evidence that the self-renewal gene NANOG regulates human tumor development. *Stem Cells* **27**(5):993-1005.

Jeter CR, Liu B, Liu X, Chen X, Liu C, Calhoun-Davis T, Repass J, Zaehres H, Shen JJ, Tang DG (2011). NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. *Oncogene* **30**(36):3833-45.

Jia L, Li H, Sun Y (2011). Induction of p21-dependent senescence by an NAE inhibitor, MLN4924, as a mechanism of growth suppression. *Neoplasia* **13**(6):561-9.

Jiang J and Hui C-c (2008). Hedgehog signaling in development and cancer. *Dev Cell* **15**, 801–812.

Jin J, Arias EE, Chen J, Harper JW, Walter JC (2006). A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. *Mol Cell* **23**(5):709-21.

Kaji K, Norrby K, Paca A, Mileikovsky M, Mohseni P, Woltjen K (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* **458**:771–775.

Kamura T, Maenaka K, Kotoshiba S, Matsumoto M, Kohda D, Conaway RC, Conaway JW, Nakayama KI (2004). VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. *Genes Dev* **18**(24):3055-65.

Kasper JS, Arai T, DeCaprio JA (2006). A novel p53-binding domain in CUL7. *Biochem Biophys Res Commun* **348**(1):132-8.

Katz EJ, Isasa M, Crosas B (2010). A new map to understand deubiquitination. *Biochem Soc Trans* **38**(1):21-8.

Kee Y, Huibregtse JM (2007). Regulation of catalytic activities of HECT ubiquitin ligases. *Biochem Biophys Res Commun* **354**:329–33.

Kerzendorfer C, Whibley A, Carpenter G, Outwin E, Chiang SC, Turner G, Schwartz C, El-Khamisy S, Raymond FL, O'Driscoll M (2010).Mutations in Cullin 4B result in a human syndrome associated with increased camptothecin-induced topoisomerase I-dependent DNA breaks. *Hum Mol Genet* **19**(7):1324-34.

Kerzendorfer C, Hart L, Colnaghi R, Carpenter G, Alcantara D, Outwin E, Carr AM, O'Driscoll M (2011).CUL4B-deficiency in humans: understanding the clinical consequences of impaired Cullin 4-RING E3 ubiquitin ligase function. *Mech Ageing Dev* **132**(8-9):366-73.

Khalfallah O, Rouleau M, Barbry P, Bardoni B, Lalli E (2009). Dax-1 knockdown in mouse embryonic stem cells induces loss of pluripotency and multilineage differentiation. *Stem Cells* **27**(7):1529-37.

Kim N, Saudemont A, Webb L, Camps M, Ruckle T, Hirsch E, Turner M, Colucci F (2007). The p110delta catalytic isoform of PI3K is a key player in NK-cell development and cytokine secretion. *Blood* **110**(9):3202-8.

Kim SS, Shago M, Kaustov L, Boutros PC, Clendening JW, Sheng Y, Trentin GA, Barsyte-Lovejoy D, Mao DY, Kay R, Jurisica I, Arrowsmith CH, Penn LZ (2007). CUL7 is a novel antiapoptotic oncogene. *Cancer Res* **67**(20):9616-22.

Kim J, Chu J, Shen X, Wang J, Orkin SH (2008). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132**:1049–1061.

Kim JB, Sebastiano V, Wu G, Araúzo-Bravo MJ, Sasse P, Gentile L, Ko K, Ruau D, Ehrich M, van den Boom D, Meyer J, Hübner K, Bernemann C, Ortmeier C, Zenke M, Fleischmann BK, Zaehres H, Schöler HR (2009). Oct4-induced pluripotency in adult neural stem cells. *Cell* **136**(3):411-9.

Kind KL, Collett RA, Harvey AJ & Thompson JG (2005). Oxygen-regulated expression of GLUT-1, GLUT-3 and VEGF in the mouse blastocyst. *Mol Reprod Dev* **70**, 37–44.

Kipreos ET, Lander LE, Wing JP, He WW, Hedgecock EM (1996). cul-1 is required for cell cycle exit in C. elegans and identifies a novel gene family. *Cell* **85**(6):829-39.

Kohl M, Wiese S, Warscheid B (2011). Cytoscape: software for visualization and analysis of biological networks. *Methods Mol Biol* **696**:291-303. Website: http://www.cytoscape.org.

Kopp JL, Ormsbee BD, Desler M, Rizzino A (2008). Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells* **26**(4):903-11.

Krupinski P, Chickarmane V, Peterson C (2011).Simulating the mammalian blastocyst-molecular and mechanical interactions pattern the embryo. *PLoS Comput Biol* **7**(5):e1001128.

Kubicek S, O'Sullivan RJ, August EM, Hickey ER, Zhang Q, Teodoro ML, Rea S, Mechtler K, Kowalski JA, Homon CA, Kelly TA, Jenuwein T (2007). Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. *Mol Cell* **25**(3):473-81.

Kuijk EW, van Mil A, Brinkhof B, Penning LC, Colenbrander B, Roelen BA (2010). PTEN and TRP53 independently suppress Nanog expression in spermatogonial stem cells. *Stem Cells Dev* **19**(7):979-88.

Kurz T, Ozlü N, Rudolf F, O'Rourke SM, Luke B, Hofmann K, Hyman AA, Bowerman B, Peter M (2005). The conserved protein DCN-1/Dcn1p is required for cullin neddylation in *C. elegans* and *S. cerevisiae*. *Nature* **435**: 1257–1261.

Larsen BD, Rampalli S, Burns LE, Brunette S, Dilworth FJ, Megeney LA (2010a). Caspase 3/caspase-activated DNase promote cell differentiation by inducing DNA strand breaks. *Proc Natl Acad Sci* **107**(9):4230-5.

Larsen BD, Megeney LA (2010b) Parole terms for a killer: directing caspase3/CAD induced DNA strand breaks to coordinate changes in gene expression. *Cell Cycle* **9**(15):2940-5.

Lawlor MA and Alessi DR (2001). PKB/Akt: a key mediator of cell proliferation, survival and insulin responses? *Journal of Cell Science* **114**, 2903-2910.

Le Bihan, T., Duewel, H.S., Figeys, D. (2003). On-line strong cation exchange micro-HPLC-ESI-MS/MS for protein identification and process optimization. *J Am Soc Mass Spectrom* **14**, 719-27.

Lee H, Sengupta N, Villagra A, Rezai-Zadeh N, Seto E (2006). Histone deacetylase 8 safeguards the human ever-shorter telomeres 1B (hEST1B) protein from ubiquitin-mediated degradation. *Mol Cell Biol.* **26**(14):5259-69.

Lee SJ, Choi JY, Sung YM, Park H, Rhim H, Kang S (2001). E3 ligase activity of RING finger proteins that interact with Hip-2, a human ubiquitin-conjugating enzyme. *FEBS Lett.* **503**(1):61-4.

Li J, Pan G, Cui K, Liu Y, Xu S, Pei D (2007). A dominant-negative form of mouse SOX2 induces trophectoderm differentiation and progressive polyploidy in mouse embryonic stem cells. *J Biol Chem* **282**(27):19481-92.

Li W, Zhou H, Abujarour R, Zhu S, Young Joo J, Lin T, Hao E, Schöler HR, Hayek A, Ding S (2009). Generation of human-induced pluripotent stem cells in the absence of exogenous Sox2. *Stem Cells* **27**(12):2992-3000.

Li W, Ding S (2010). Small molecules that modulate embryonic stem cell fate and somatic cell reprogramming. *Trends Pharmacol Sci* **31**(1):36-45.

Li W, Jiang K, Ding S (2012). Concise review: A chemical approach to control cell fate and function. *Stem Cells* **30**(1):61-8.

Li X, Sun L, Jin Y (2008). Identification of karyopherin-α 2 as an Oct4 associated protein. *J Genet Genomics* **35**:723–728.

Li X, Lu D, He F, Zhou H, Liu Q, Wang Y, Shao C, Gong Y. (2011). Cullin 4B protein ubiquitin ligase targets peroxiredoxin III for degradation. *J Biol Chem*. **286**(37):32344-54.

Licklider, L.J., Thoreen, C.C., Peng, J., Gygi, S.P. (2002). Automation of nanoscale microcapillary liquid chromatography-tandem mass spectrometry with a vented column. *Anal Chem* **74**, 3076-83.

Lin JJ, Milhollen MA, Smith PG, Narayanan U, Dutta A (2010). NEDD8-targeting drug MLN4924 elicits DNA rereplication by stabilizing Cdt1 in S phase, triggering checkpoint activation, apoptosis, and senescence in cancer cells. *Cancer Res* **15**; 70(24):10310-20.

Lin T, Chao C, Saito S, Mazur SJ, Murphy ME, Appella E, Xu Y (2005). p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. *Nat Cell Biol* **7**(2):165-71.

Liang J, Wan M, Zhang Y, Gu P, Xin H, Jung SY, Qin J, Wong J, Cooney AJ, Liu D, Songyang Z (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nat Cell Biol* **10**:731–739.

Liedtke S, Enczmann J, Waclawczyk S, Wernet P, Kögler G (2007). Oct4 and its pseudogenes confuse stem cell research. *Cell Stem Cell* **1**(4):364-6.

Lochner A, Moolman JA (2006). The many faces of H89: a review. *Cardiovasc Drug Rev* **24**(3-4):261-74.

Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet* **38**:431–440.

Lowery DM, Clauser KR, Hjerrild M, Lim D, Alexander J, Kishi K, Ong SE, Gammeltoft S, Carr SA, Yaffe MB (2007). Proteomic screen defines the Polo-box domain interactome and identifies Rock2 as a Plk1 substrate. *EMBO J*. **26**(9):2262-73.

Lowry WE, Plath K (2008). The many ways to make an iPS cell. *Nat Biotechnol* **26**(11):1246-8.

Lu Z, Hunter T (2009). Degradation of Activated Protein Kinases by Ubiquitination. *Annu Rev Biochem* **78**: 435-475.

Luke-Glaser S, Roy M, Larsen B, Le Bihan T, Metalnikov P, Tyers M, Peter M, Pintard L. (2007). CIF-1, a shared subunit of the COP9/signalosome and eukaryotic initiation factor 3 complexes, regulates MEL-26 levels in the Caenorhabditis elegans embryo. *Mol Cell Biol* **27**(12):4526-40.

Lyssiotis CA, Lairson LL, Boitano AE, Wurdak H, Zhu S, Schultz PG (2011). Chemical control of stem cell fate and developmental potential. *Angew Chem Int Ed Engl* **50**(1):200-42.

Ma T, Wang Z, Guo Y, Pei D (2009). The C-terminal pentapeptide of Nanog tryptophan repeat domain interacts with Nac1 and regulates stem cell proliferation but not pluripotency. *J Biol Chem* **284**(24):16071-81.
Maherali N, Hochedlinger K (2009). Tgfbeta signal inhibition cooperates in the induction of iPSCs and replaces Sox2 and cMyc. *Curr Biol* **19**(20):1718-23.

Mahrour N, Redwine WB, Florens L, Swanson SK, Martin-Brown S, Bradford WD, Staehling-Hampton K, Washburn MP, Conaway RC, Conaway JW (2008). Characterization of Cullin-box sequences that direct recruitment of Cul2-Rbx1 and Cul5-Rbx2 modules to Elongin BC-based ubiquitin ligases. *J Biol Chem* **283**(12):8005-13.

Mak AB, Ni Z, Hewel JA, Chen GI, Zhong G, Karamboulas K, Blakely K, Smiley S, Marcon E, Roudeva D, Li J, Olsen JB, Wan C, Punna T, Isserlin R, Chetyrkin S, Gingras AC, Emili A, Greenblatt J, Moffat J (2010). A lentiviral functional proteomics approach identifies chromatin remodeling complexes important for the induction of pluripotency. *Mol Cell Proteomics* **9**(5):811-23.

Malki S, Nef S, Notarnicola C, Thevenet L, Gasca S, Méjean C, Berta P, Poulat F, Boizet-Bonhoure B (2005). Prostaglandin D2 induces nuclear import of the sex-determining factor SOX9 via its cAMP-PKA phosphorylation. *EMBO J* **24**, 1798-809.

Mallanna SK, Ormsbee BD, Iacovino M, Gilmore JM, Cox JL, Kyba M, Washburn MP, Rizzino A (2010). Proteomic analysis of Sox2-associated proteins during early stages of mouse embryonic stem cell differentiation identifies Sox21 as a novel regulator of stem cell fate. *Stem Cell* **28**:1715–1727.

Mallanna SK, Rizzino A (2012). Systems biology provides new insights into the molecular mechanisms that control the fate of embryonic stem cells. *J Cell Physiol* **227**(1):27-34.

Marmorstein LY, Kinev AV, Chan GK, Bochar DA, Beniya H, Epstein JA, Yen TJ, Shiekhattar R (2001).A human BRCA2 complex containing a structural DNA binding component influences cell cycle progression. *Cell* **104**(2):247-57.

Maruyama M, Ichisaka T, Nakagawa M, Yamanaka S (2005). Differential roles for Sox15 and Sox2 in transcriptional control in mouse embryonic stem cells. *J Biol Chem* **280**(26):24371-9.

Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov AA, Ko MS, Niwa H (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* **9**(6):625-35.

Matheny SA, White MA (2006). Ras-sensitive IMP modulation of the Raf/MEK/ERK cascade through KSR1. *Methods Enzymol* **407**:237-47.

Matoba R, Niwa H, Masui S, Ohtsuka S, Carter MG, Sharov AA, Ko MS (2006). Dissecting Oct3/4-regulated gene networks in embryonic stem cells by expression profiling. *PLoS ONE* **1**:e26.

Matsuda T, Nakamura T, Nakao K, Arai T, Katsuki M, Heike T, Yokota T (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *EMBO J* **18**(15):4261-9.

Meng HM, Zheng P, Wang XY, Liu C, Sui HM, Wu SJ, Zhou J, Ding YQ, Li JM (2010). Overexpression of nanog predicts tumor progression and poor prognosis in colorectal cancer. *Cancer Biol Ther* **9**(4).

Meraldi P, Draviam VM, Sorger PK (2004). Timing and checkpoints in the regulation of mitotic progression. *Dev Cell* **7**(1):45-60.

Mikenberg I, Widera D, Kaus A, Kaltschmidt B, Kaltschmidt C (2007). Transcription Factor NF-κB Is Transported to the Nucleus via Cytoplasmic Dynein/Dynactin Motor Complex in Hippocampal Neurons. *PLoS ONE* 2(7): e589.

Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, Bernstein BE, Jaenisch R, Lander ES, Meissner A (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**(7200):49-55.

Milhollen MA, Traore T, Adams-Duffy J, Thomas MP, Berger AJ, Dang L, Dick LR, Garnsey JJ, Koenig E, Langston SP, Manfredi M, Narayanan U, Rolfe M, Staudt LM, Soucy TA, Yu J, Zhang J, Bolen JB, Smith PG (2009). MLN4924, a NEDD8-activating enzyme inhibitor, is active in diffuse large B-cell lymphoma models: rationale for treatment of NF-{kappa}B-dependent lymphoma. *Blood* **16**(9):1515-23.

Milhollen MA, Narayanan U, Soucy TA, Veiby PO, Smith PG, Amidon B (2011). Inhibition of NEDD8-activating enzyme induces rereplication and apoptosis in human tumor cells consistent with deregulating CDT1 turnover. *Cancer* **15**; 71(8):3042-51.

Milyavsky M, Gan OI, Trottier M, Komosa M, Tabach O, Notta F, Lechman E, Hermans KG, Eppert K, Konovalova Z, Ornatsky O, Domany E, Meyn MS, Dick JE (2010) A distinctive DNA damage response in human hematopoietic stem cells reveals an apoptosis-independent role for p53 in self-renewal. *Cell Stem Cell* **7**(2):186-97.

Mitalipov S, Wolf D (2009). Totipotency, pluripotency and nuclear reprogramming. *Adv Biochem Eng Biotechnol* **114**:185-99.

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**(5):631-42.

Moon JH, Kwon S, Jun EK, Kim A, Whang KY, Kim H, Oh S, Yoon BS, You S (2011). Nanog-induced dedifferentiation of p53-deficient mouse astrocytes into brain cancer stem-like cells. *Biochem Biophys Res Commun* **412**(1):175-81.

Moore KA, Lemischka IR (2006). Stem cells and their niches. *Science* **311**:1880-5.

Moretto-Zita M, Jin H, Shen Z, Zhao T, Briggs SP, Xu Y (2010). Phosphorylation stabilizes Nanog by promoting its interaction with Pin1. *Proc Natl Acad Sci U S A*. **107**(30):13312-7.

Mullin NP, Yates A, Rowe AJ, Nijmeijer B, Colby D, Barlow PN, Walkinshaw MD, Chambers I. (2008). The pluripotency rheostat Nanog functions as a dimer. *Biochem J*. **411**(2):227-31.

Nakagawa T, Xiong Y (2011). Chromatin regulation by CRL4 E3 ubiquitin ligases: CUL4B targets WDR5 ubiquitylation in the nucleus. *Cell Cycle* **10**(24):4197-8.

Nichols J, Evans EP, Smith AG (1990). Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. *Development* **110**(4):1341-8.

Nichols, J (2001). Introducing embryonic stem cells. *Curr Biol* **11**, R503–R505.

Nichols J, Ying QL (2006). Derivation and propagation of embryonic stem cells in serum- and feeder-free culture. Methods Mol Biol 329:91-8.

Nichols J, Smith A (2009). Naive and primed pluripotent states. *Cell Stem Cell* **4**(6):487-92.

NIH report on stem cells (2001). Stem Cells: Scientific Progress and Future Research Directions. Department of Health and Human Services.

Niwa H, Burdon T, Chambers I, Smith A (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* **12**(13):2048-60.

Niwa H, Miyazaki J, Smith AG (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**:372–376.

Niwa H (2007). How is pluripotency determined and maintained? (2007) *Development* **34**(4):635-46.

Ochocka AM, Kampanis P, Nicol S, Allende-Vega N, Cox M, Marcar L, Milne D, Fuller-Pace F, Meek D (2009).FKBP25, a novel regulator of the p53 pathway, induces the degradation of MDM2 and activation of p53. *FEBS Lett* **583**(4):621-6.

Ohnishi T, Yamashita A, Kashima I, Schell T, Anders KR, Grimson A, Hachiya T, Hentze MW, Anderson P, Ohno S (2003). Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. *Mol Cell.* **12**(5):1187-200.

Okita K, Ichisaka T, Yamanaka S (2007). Generation of germline-competent induced pluripotent stem cells. *Nature* **448**(7151):313-7.

Okita K, Nakagawa M, Hyenjong H, Ichisaka T, Yamanaka S (2008). Generation of mouse induced pluripotent stem cells without viral vectors. *Science* **322**(5903):949-53.

Olma MH, Roy M, Le Bihan T, Sumara I, Maerki S, Larsen B, Quadroni M, Peter M, Tyers M, Pintard L (2009). An interaction network of the mammalian COP9 signalosome identifies Dda1 as a core subunit of multiple Cul4-based E3 ligases. *J Cell Sci* **122**(Pt 7):1035-44.

Papapetrou EP, Tomishima MJ, Chambers SM, Mica Y, Reed E, Menon J, Tabar V, Mo Q, Studer L, Sadelain M (2009). Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proc Natl Acad Sci U S A* **106**(31):12759-64.

Pardo M, Lang B, Yu L, Prosser H, Bradley A, Babu MM, Choudhary J (2010). An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell Stem Cell* **6**(4):382-95.

Pease S, Williams RL (1990). Formation of germ-line chimeras from embryonic stem cells maintained with recombinant leukemia inhibitory factor. *Exp Cell Res* **190**(2):209-11.

Peltier J, Schaffer DV (2010). Systems biology approaches to understanding stem cell fate choice. *IET Syst Biol* **4**(1):1-11.

Petroski MD, Deshaies RJ (2005). Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol* **6**(1):9-20.

Piestun D, Kochupurakkal B S, Jacob-Hirsch J, Zeligson S, Koudritsky M, Domany E, Amariglio N, Rechavi G, Givol D (2006). Nanog transforms NIH3T3 cells and targets cell-type restricted genes. *Biochem Biophys Res Commun* **343**, 279-285.

Pintard L, Willis JH, Willems A, Johnson JL, Srayko M, Kurz T, Glaser S, Mains PE, Tyers M, Bowerman B, Peter M (2003).The BTB protein MEL-26 is a substrate-specific adaptor of the CUL-3 ubiquitin-ligase. *Nature* **425**(6955):311-6.

Pintard L, Willems A, Peter M (2004). Cullin-based ubiquitin ligases: Cul3-BTB complexes join the family. *EMBO J* **23**(8):1681-7.

Pires IM, Bencokova Z, Milani M, Folkes LK, Li JL, Stratford MR, Harris AL, Hammond EM (2010). Effects of acute versus chronic hypoxia on DNA damage responses and genomic instability. *Cancer Res* **70**(3):925-35.

Ramakrishna S, Suresh B, Lim KH, Cha BH, Lee SH, Kim KS, Baek KH (2011). PEST motif sequence regulating human NANOG for proteasomal degradation. *Stem Cells Dev* **20**(9):1511-9.

Reyes-Turcu FE, Ventii KH, Wilkinson KD (2009). Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annu Rev Biochem* **78**:363-97.

Rozenblatt-Rosen O, Hughes CM, Nannepaga SJ, Shanmugam KS, Copeland TD, Guszczynski T, Resau JH, Meyerson M (2005) The parafibromin tumor suppressor protein is part of a human Paf1 complex. *Mol Cell Biol.* **25**(2):612-20.

Ruggero D, Grisendi S, Piazza F, Rego E, Mari F, Rao PH, Cordon-Cardo C, Pandolfi PP. (2003). Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. *Science* **299**(5604):259-62.

Sadek H, Hannack B, Choe E, Wang J, Latif S, Garry MG, Garry DJ, Longgood J, Frantz DE, Olson EN, Hsieh J, Schneider JW (2008). Cardiogenic small molecules that enhance myocardial repair by stem cells. *Proc Natl Acad Sci U S A* **105**(16):6063-8.

Sarikas A, Xu X, Field LJ, Pan ZQ (2008).The cullin7 E3 ubiquitin ligase: a novel player in growth control. *Cell Cycle* **7**(20):3154-61.

Sarkisian, C J, Master S R, Huber L J, Ha S I, Chodosh L A (2001). Analysis of murine Brca2 reveals conservation of protein-protein interactions but differences in nuclear localization signals. *J Biol Chem* **276** (40): 37640–8.

Satijn D P, Gunster M J, van der Vlag J, Hamer K M, Schul W, Alkema M J, Saurin A J, Freemont P S, van Driel R, Otte A P (1997). RING1 is associated with the polycomb group protein complex and acts as a transcriptional repressor. *Mol Cell Biol* **17**(7): 4105–13.

Satijn D P and Otte A P (1999). RING1 interacts with multiple Polycomb-group proteins and displays tumorigenic activity. *Mol Cell Biol* **19**(1): 57–68.

Savatier, P., Lapillonne, H., van Grunsven, L. A., Rudkin, B. B., and Samarut, J. (1996). Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor up-regulates D-type cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. *Oncogene* **12**, 309–322.

Saxe JP, Tomilin A, Schöler HR, Plath K, Huang J (2009). Post-Translational Regulation of Oct4 Transcriptional Activity. *PLoS ONE* **4**(2): e4467.

Schägger H, von Jagow G (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem* **199**(2):223-31.

Schugar RC, Robbins PD, Deasy BM (2008). Small molecules in stem cell self-renewal and differentiation. *Gene Ther* **15**(2):126-35.

Seamon KB, Daly JW (1981). Forskolin: a unique diterpene activator of cyclic AMP-generating systems. *J Cyclic Nucleotide Res* **7**(4):201-24.

Seki Y, Kurisaki A,Watanabe-Susaki K, Nakajima Y, Nakanishi M, Arai Y, Shiota K, Sugino H, Asashima M (2010). TIF1beta regulates the pluripotency of embryonic stem cells in a phosphorylation-dependent manner. *Proc Natl Acad Sci USA* **107**:10926–10931.

Semenza GL (2011). Regulation of Metabolism by Hypoxia-Inducible Factor 1. *Cold Spring Harb Symp Quant Biol* Jul 22.

Semenza GL (2011).Oxygen sensing, homeostasis, and disease. *N Engl J Med* **365**(6):537-47.

Shi Y, Desponts C, Do JT, Hahm HS, Scholer HR, Ding S (2008). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* **3**:568–574.

Shimada H, Hashimoto Y, Nakada A, Shigeno K, Nakamura T (2012). Accelerated generation of human induced pluripotent stem cells with retroviral transduction and chemical inhibitors under physiological hypoxia. *Biochem Biophys Res Commun* **417**(2):659-64.

Silva J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, Smith A (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol* **6**(10):e253.

Singh AM, Hamazaki T, Hankowski KE, Terada N(2007). A heterogeneous expression pattern for Nanog in embryonic stem cells. *Stem Cells* **25**:2534–2542.

Smith AG, Heath JK, Donaldson DD, Wong GG, Moreau J, Stahl M, Rogers D (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**(6200):688-90.

Smith AG (1992). Mouse embryo stem cells: their identification, propagation and manipulation. *Semin Cell Biol* **3**(6):385-99.

Smith AG (2001). Embryo-derived stem cells: of mice and men. *Annu Rev Cell Dev Biol* **17**:435-62.

Smith A (2001). Embryonic stem cells. In Stem Cell Biology (Marshak, DR *et al.*, eds), pp. 205–230, Cold Spring Harbor Laboratory Press.

Solozobova V, Rolletschek A, Blattner C (2009). Nuclear accumulation and activation of p53 in embryonic stem cells after DNA damage. *BMC Cell Biol* **10**:46.

Soucy TA, Smith PG, Milhollen MA, Berger AJ, Gavin JM, Adhikari S, Brownell JE, Burke KE, Cardin DP, Critchley S (2009). An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458**: 732–736.

Stadtfeld M, Nagaya M, Utikal J, Weir G, Hochedlinger K (2008). Induced pluripotent stem cells generated without viral integration. *Science* **322**(5903):945-9.

Suda Y, Suzuki M, Ikawa Y, Aizawa S (1987). Mouse embryonic stem cells exhibit indefinite proliferative potential. *J Cell Physiol* **133**(1):197-201.
Sun Y (2003). Targeting E3 ubiquitin ligases for cancer therapy. *Cancer Biol Ther* **2**:623–9.

Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, Jensen LJ, von Mering C (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* **39** (Database issue) D561-8. STRING v 9.0 Website: http://string-db.org.

Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr Biol* **11**(19):1553-8.

Takahashi K, Mitsui K, Yamanaka S (2003). Role of ERas in promoting tumour-like properties in mouse embryonic stem cells. *Nature* **423**(6939):541-5.

Takahashi K, Yamanaka S (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**(4):663-76.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**(5):861-72.

Takao Y, Yokota T, Koide H (2006). β-catenin up-regulates Nanog expression through interaction with Oct-3/4 in embryonic stem cells. *Biochem Biophys Res Commun* **353**:699–705.

Thrower JS, Hoffman L, Rechsteiner M, Pickart CM (2000). Recognition of the polyubiquitin proteolytic signal. *EMBO J* **19**(1):94-102.

Tommer Ravid & Mark Hochstrasser (2008). Diversity of degradation signals in the ubiquitin–proteasome system. *Nature Reviews Molecular Cell Biology* **9**, 679-689.

Toyooka Y, Tsunekawa N, Akasu R, Noce T (2003). Embryonic stem cells can form germ cells in vitro. *Proc Natl Acad Sci* **100**(20):11457-62.

Toyooka Y, Shimosato D, Murakami K, Takahashi K, Niwa H (2008)Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* **135**:909–918.

Tripathi R, Sastry KS, Kota SK, Srinivas UK (2005). Cloning and characterization of mouse cullin4B/E3 ubiquitin ligase. *J Biosci* **30**(3):329-37.

Tripathi R, Kota SK, Srinivas UK (2007). Cullin4B/E3-ubiquitin ligase negatively regulates β-catenin. *J Biosci* **32**:1133–1138.

Tsai LL, Yu CC, Chang YC, Yu CH, Chou MY (2011). Markedly increased Oct4 and Nanog expression correlates with cisplatin resistance in oral squamous cell carcinoma. *J Oral Pathol Med* **40**(8):621-8.

Turksen K (2002). Embryonic stem cells: methods and protocols. *Humana Press* **185**; Volume 2002.

Van den Berg D L, Zhang W, Yates A, Engelen E, Takacs K, Bezstarosti K, Demmers J, Chambers I, and Poot R A (2008). Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. *Mol Cell Biol* **28**, 5986-5995.

Van den Berg DL, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, Chambers I, Poot RA (2010). An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* **6**(4):369-81.

van Mourik MS, Macklon NS, Heijnen CJ (2008). Embryonic implantation: cytokines, adhesion molecules, and immune cells in establishing an implantation environment. *J Leukoc Biol* **85**(1):4-19.

Vickaryous, M. K., and Hall, B. K. (2006) Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest, *Biol Rev Cambridge Philos Soc* **81**, 425–455.

Walsh DA, Glass DB (1991). Utilization of the inhibitor protein of adenosine cyclic monophosphate-dependent protein kinase, and peptides derived from it, as tools to study adenosine cyclic monophosphate-mediated cellular processes. *Methods Enzymol* **201**:304-16.

Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature 431(7010):873-8.

Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH (2006). A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**(7117):364-8.

Wang J, Levasseur DN, Orkin SH (2008). Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. *Proc Natl Acad Sci* **105**(17):6326-31.

Wang J, Orkin SH (2008). A protein roadmap to pluripotency and faithful reprogramming. *Cells Tissues Organs* **188**(1-2): 23-30.

Wang Y, Cortez D, Yazdi P, Neff N, Elledge SJ, Qin J (2000). BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev* **14**(8):927-39.

Wang Y, Chen Y, Chen M, Xu W (2006). AKAPs competing peptide HT31 disrupts the inhibitory effect of PKA on RhoA activity. *Oncol Rep***16**(4):755-61.

Warashina M, Min KH, Kuwabara T, Huynh A, Gage FH, Schultz PG, Ding S (2006). A synthetic small molecule that induces neuronal differentiation of adult hippocampal neural progenitor cells. *Angew Chem Int Ed Engl* **45**(4):591-3.

Washburn MP, Koller A, Oshiro G, Ulaszek RR, Plouffe D, Deciu C, Winzeler E, Yates JR 3rd (2003). Protein pathway and complex clustering of correlated mRNA and protein expression analyses in Saccharomyces cerevisiae. *Proc Natl Acad Sci USA* **100**(6):3107-12.

White J, Dalton S (2005). Cell cycle control of embryonic stem cells. *Stem Cell Rev* **1**(2):131-8.

Whiteside ST, Goodbourn S (1993) Signal transduction and nuclear targeting: regulation of transcription factor activity by subcellular localisation. *Journal of cell science* **104** (4): 949–55.

Wilkinson KD (1997). Regulation of ubiquitin-dependent processes by deubiquitinating enzymes. *FASEB J* **11**(14):1245-56.

Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**(6200):684-7.

Willems AR, Schwab M, Tyers M (2004). A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochim Biophys Acta* **1695**(1-3):133-70.

Winston JT, Koepp DM, Zhu C, Elledge SJ, Harper JW (1999). A family of mammalian F-box proteins. *Curr Biol* **9**(20):1180-2.

Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, Hämäläinen R, Cowling R, Wang W, Liu P, Gertsenstein M, Kaji K, Sung HK, Nagy A (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* **458**(7239):766-70.

Wray J, Kalkan T, Smith AG (2010). The ground state of pluripotency. *Biochem Soc Trans* **38**:1027–1032.

Wu Q, Chen X, Zhang J, Loh YH, Low TY, Zhang W, Zhang W, Sze SK, Lim B, Ng HH (2006). Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. J Biol Chem **281**(34):24090-4.

Wu MY and Hill CS (2009). Tgf- superfamily signaling in embryonic development and homeostasis, *Dev Cell* **16**, 329–343.

Xu L, Wei Y, Reboul J, Vaglio P, Shin TH, Vidal M, Elledge SJ, Harper JW (2003). BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. *Nature* **425**(6955):316-21.

Xu Y (2005). A new role for p53 in maintaining genetic stability in embryonic stem cells. *Cell Cycle* **4**(3):363-4.

Yamamizu K, Fujihara M, Tachibana M, Katayama S, Takahashi A, Hara E, Imai H, Shinkai Y, Yamashita JK. (2012). Protein Kinase A Determines Timing of Early Differentiation through Epigenetic Regulation with G9a. *Cell Stem Cell* **10**(6):759-70.

Yamasaki C, Tashiro S, Nishito Y, Sueda T, Igarashi K. Dynamic cytoplasmic anchoring of the transcription factor Bach1 by intracellular hyaluronic acid binding protein IHABP (2005). *J Biochem* **137**(3):287-96.

Yang, X., Downes, M., Yu, R. T., Bookout, A. L., He, W., Straume, M., Mangelsdorf, D. J. and Evans, R. M. (2006) Nuclear receptor expression links the circadian clock to metabolism *Cell* **126**, 801-10.

Yates A, Chambers I (2005). The homeodomain protein Nanog and pluripotency in mouse embryonic stem cells. *Biochem Soc Trans* **33**(Pt 6):1518-21.

Ying QL, Nichols J, Chambers I, Smith A (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115**(3):281-92.

Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A (2008). The ground state of embryonic stem cell self-renewal. *Nature* **453**(7194): 519-23.

Yoshida K, Miki Y (2004). Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. *Cancer science* **95** (11): 866–71.

Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318** (5858):1917-20.

Yuri S, Fujimura S, Nimura K, Takeda N, Toyooka Y, Fujimura Y, Aburatani H, Ura K, Koseki H, Niwa H, Nishinakamura R (2009). Sall4 is essential for stabilization, but not for pluripotency, of embryonic stem cells by repressing aberrant trophectoderm gene expression. *Stem Cells* **27**(4):796-805.

Zbinden M, Duquet A, Lorente-Trigos A, Ngwabyt SN, Borges I, Ruiz i Altaba A (2010). NANOG regulates glioma stem cells and is essential in vivo acting in a cross-functional network with GLI1 and p53. *EMBO J* **29**(15):2659-74.

Zhang DD, Lo SC, Cross JV, Templeton DJ, Hannink M (2004). Keap1 is a redox-regulated substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex. *Mol Cell Biol* **24**(24):10941-53.

Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY, Soh BS, Lou Y, Yang J, Ma Y, Chai L, Ng HH, Lufkin T, Robson P, Lim B (2006). Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol* **8**(10):1114-23.

Zhang X, Zhang J, Wang T, Esteban MA, Pei D (2008). Esrrb Activates Oct4 Transcription and Sustains Selfrenewal and Pluripotency in Embryonic Stem Cells. *J Biol Chem* **283**: 35825-35833.

Zhang X, Neganova I, Przyborski S, Yang C, Cooke M, Atkinson SP, Anyfantis G, Fenyk S, Keith WN, Hoare SF, Hughes O, Strachan T, Stojkovic M, Hinds PW, Armstrong L, Lako M (2009). A role for NANOG in G1 to S transition in human embryonic stem cells through direct binding of CDK6 and CDC25A. *J Cell Biol* **184**(1):67-82.

Zhao HY, Zhang YJ, Dai H, Zhang Y, Shen YF (2011). CARM1 mediates modulation of Sox2. *PLoS One* **6**(10):e27026.

Zheng N, Schulman BA, Song L, Miller JJ, Jeffrey PD, Wang P, Chu C, Koepp DM, Elledge SJ, Pagano M, Conaway RC, Conaway JW, Harper JW, Pavletich NP (2002). Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**(6882):703-9.

Zou Y, Liu Q, Chen B, Zhang X, Guo C, Zhou H, Li J, Gao G, Guo Y, Yan C, Wei J, Shao C, Gong Y (2007). Mutation in CUL4B, which encodes a member of cullin-RING ubiquitin ligase complex, causes X-linked mental retardation. *Am J Hum Genet.* **80**(3):561-6.

Zou Y, Mi J, Cui J, Lu D, Zhang X, Guo C, Gao G, Liu Q, Chen B, Shao C, Gong Y (2009). Characterization of nuclear localization signal in the N terminus of CUL4B and its essential role in cyclin E degradation and cell cycle progression. *J Biol Chem* **284**(48):33320-32.

Zhu B, Zheng Y, Pham AD, Mandal SS, Erdjument-Bromage H, Tempst P, Reinberg D (2005). Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol Cell.* **20**(4):601-11.